



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

도시계획학 석사학위논문

대중교통 통행배정을 위한
링크기반 확률적 알고리즘 연구:
도시철도 통행배정을 대상으로

A Link-based Stochastic Algorithm for Transit Assignment:
Results from the Urban Railway System in Seoul

2012년 8월

서울대학교 환경대학원
환경계획학과 교통관리전공
정 동 재

대중교통 통행배정을 위한
링크기반 확률적 알고리즘 연구:
도시철도 통행배정을 대상으로

지도교수 장 수 은

이 논문을 도시계획학 석사학위 논문으로 제출함
2012년 4월

서울대학교 환경대학원
환경계획학과 교통관리전공
정 동 재

정동재의 도시계획학 석사 학위논문을 인준함
2012년 6월

위 원 장 _____(인)

부위원장 _____(인)

위 원 _____(인)

국 문 초 록

대중교통 통행배정 알고리즘은 전통적인 승용차 통행배정 알고리즘의 분류 방법과 마찬가지로 통행자의 인지오차 고려여부에 따라 결정적 통행배정 알고리즘과 확률적 통행배정 알고리즘으로 분류할 수 있다. 또한 분석단위가 링크인지 경로인지에 따라 다시 링크기반 통행배정 알고리즘과 경로기반 통행배정 알고리즘으로 나뉜다.

현재 표준적으로 사용하고 있는 대중교통 통행배정 알고리즘은 링크기반 결정적 통행배정 알고리즘과 경로기반 확률적 통행배정 알고리즘이다. 그런데 두 알고리즘은 각각 통행자의 경로선택에 대한 ‘결정론적’인 가정과 분석단위가 ‘경로’라는 측면에서 결함이 나타난다. 본 연구는 위의 두 알고리즘의 한계점을 극복하기 위한 대안으로 링크기반 확률적 통행배정 알고리즘을 제안한다. 특히 링크기반 확률적 통행배정 알고리즘 중에서도 가장 대표적인 Dial 알고리즘을 통해서 그 적합성을 살펴본다. 이에 앞서, 승용차 통행배정을 목적으로 고안된 Dial 알고리즘을 대중교통 통행배정에 적용하기 위하여 대중교통의 특징적인 통행 특성인 환승통행을 반영할 수 있는 네트워크 표현방식을 구축한다. 또한 전통적 Dial 알고리즘은 개별 출발지-도착지 쌍 별로 경로분석을 함에 있어 취약점을 갖는데, 이러한 한계점을 해결하기 위하여 Dial 알고리즘을 수정 및 보완한다. 파라미터 θ 의 정산은 Robillard(1974)의 정산 방법을 준용하여 실시한다.

링크기반 결정적 통행배정 알고리즘, 경로기반 확률적 통행배정 알고리즘, 링크기반 확률적 통행배정 알고리즘의 적합성을 살펴보기 위하여 각 유형별 대표적인 알고리즘인 최적전략 알고리즘, 로짓기반 알고리즘, 본 연구를 통해 수정 및 보완된 Dial 알고리즘을 예제 네트워크와 실제 수도권 도시철도 네트워크에 적용해 본다. 이를 통해 링크기반 확

를적 통행배정 알고리즘 적용에 대한 당위성을 논의한다.

먼저, 링크기반 결정적 통행배정 알고리즘인 최적전략 알고리즘은 대중교통 이용자의 통행행태를 설명하기 위하여 승차시간, 대기시간 인자, 대기시간 가중치, 승차시간 가중치, 접근시간 가중치 등 이용자 중심적인 다수의 변수와 파라미터를 사용하고 있지만 전량배정방식이 전제하는 결정론적인 가정때문에 이용자의 통행행태를 설명하는데 한계가 있다. 특히 경로선택 확률이 노선의 운행빈도 비율의 함수로서 다른 서비스 수준은 고려하지 못하기 때문에 배차간격이 비슷하고, 이용자의 경로선택에 영향을 미치는 요소가 많은 복잡한 수도권 도시철도 네트워크 분석에 부적합한 것으로 판단된다.

경로기반 확률적 통행배정 알고리즘은 통행자들이 같은 경로를 사용하더라도 그 통행시간을 다르게 인지한다고 가정함으로써 링크기반 결정적 통행배정 알고리즘의 ‘결정론적’인 가정을 완화한다. 그러나 이용자의 경로선택에 영향을 미치는 속성변수가 상당히 많아서 효용함수의 추정은 시간 및 비용이 많이 소요되는 어려운 작업이며, 그 결과에 따라 통행배정 결과가 민감하게 달라지는 한계점을 살펴볼 수 있다. 아울러, 경로기반 확률적 통행배정 알고리즘은 분석단위가 ‘경로’이기 때문에 출발지-도착지 사이에 존재하는 모든 경로를 나열(enumeration)해야 하는 문제가 발생한다. 많은 수의 링크와 노드로 구성된 대규모 네트워크에서는 출발지와 도착지 간에 무수히 많은 경로가 존재하기 때문에 모든 경로를 나열하는 것은 사실상 불가능하다. 또한 통행자들이 경로를 선택할 때 고려하지 않는 경로에도 선택확률을 배분함으로써 그만큼의 오차가 발생할 수 있다.

반면에, 링크기반 확률적 통행배정 알고리즘인 Dial 알고리즘은 링크기반 결정적 통행배정 알고리즘과 경로기반 확률적 통행배정 알고리즘의 내재적 결함을 극복한다. 파라미터 θ 를 통해 인지도차를 반영하고

단일 경로에 전량배정하는 문제점을 완화한다. 또한 비합리적인 경로를 효율적 경로의 정의를 통해 사전에 제거하고, 링크와 노드 정보만을 이용하여 링크의 이용확률을 계산함으로써 경로나열 문제를 해결한다. 게다가 많은 변수를 요구하지 않으면서도 실제 대중교통 통행특성을 유연하게 묘사하고 이례적인 통행특성이 나타나는 구간에서도 범용적으로 적용할 수 있는 장점을 갖는다.

.....

주요어 : 대중교통 통행배정, 링크기반 결정적 통행배정, 링크기반 확률적 통행배정, 경로기반 확률적 통행배정, Dial 알고리즘

학 번 : 2010-23895

목 차

I . 서론	1
1. 연구의 배경 및 목적	1
2. 연구의 대상 및 범위	3
3. 연구의 구성 및 방법	4
II . 선행 연구 고찰	6
1. 대중교통 통행배정 알고리즘의 개요	6
2. 링크기반 결정적 알고리즘	10
3. 경로기반 확률적 알고리즘	21
4. 선행연구의 시사점	25
III . 링크기반 확률적 대중교통 통행배정 알고리즘	28
1. 개요	28
2. Dial 알고리즘	30
1) Dial 알고리즘의 개요	30
2) 인지도차와 파라미터 θ	33
3) 로짓기반 알고리즘	35
4) 효율적 경로	38
3. Dial 알고리즘의 수정	46
1) 네트워크 표현	46

2) 출발지-도착지 별 링크이용 확률	50
4. 파라미터 θ 의 정산	53

IV. 알고리즘의 성과도 평가 56

1. 개요	56
2. 예제 네트워크 기반 평가	56
1) 예제 네트워크 구축	56
2) 최적전략 알고리즘 적용	57
3) 경로기반 확률적 알고리즘 적용	58
4) 링크기반 확률적 알고리즘 적용	62
5) 소결론	65
3. 실제 네트워크 기반 평가	67
1) 분석지점 및 자료 선정	67
2) 최적전략 알고리즘 적용	70
3) 경로기반 확률적 알고리즘 적용	77
4) 링크기반 확률적 알고리즘 적용	81
가) 네트워크 구축	81
나) 최단경로 및 최단거리 $p(i)$ 산정	84
다) 수정된 Dial 알고리즘 구축	88
라) 효율적 경로 및 선택대안집합	92
마) 알고리즘의 정산	94
바) 알고리즘의 적용	101

사) 알고리즘의 검증	104
5) 소결론	106
V. 결론	109
1. 연구의 요약 및 종합적 해석	109
2. 연구의 한계 및 향후 과제	114
■ 참고문헌	119

표 목 차

<표 II-1> 통행배정 알고리즘의 분류	7
<표 II-2> 연구의 범위	9
<표 II-3> 각 출발역에서 도착역 B까지의 전략, 대기시간 및 이용확률	12
<표 II-4> 대기시간 가중치가 2일 때의 최적전략	18
<표 II-5> 대기시간 가중치가 100일 때의 최적전략	18
<표 II-6> 효용함수(또는 일반화비용) 변수 및 파라미터 추정 결과	23
<표 III-1> 서울대입구-금호 간 전체 경로 및 효율적 경로	39
<표 III-2> 신대방-서울역 간 전체 경로 및 효율적 경로	44
<표 IV-1> 출발지(A)-도착지(B) 간 경로 및 선택확률	58
<표 IV-2> 파라미터 추정결과	59
<표 IV-3> 출발지(A)-도착지(B) 간 선택가능 경로집합 및 경로속성	60
<표 IV-4> 경로기반 알고리즘의 경로선택확률	61
<표 IV-5> θ 값에 따른 경로선택 확률(대기시간 가중치=1)	64
<표 IV-6> θ 값에 따른 경로선택 확률(대기시간 가중치=1.5)	65
<표 IV-7> 분석지점(O-D) 및 실제 이용된 경로 정보	68
<표 IV-8> 지하철 평균표정속도 및 평균환승시간	69
<표 IV-9> 지하철 노선 별 배차간격	69
<표 IV-10> 수도권 지역의 1시간 통행량의 지속시간과 각 시간대별 집중률	71
<표 IV-11> 노선 별 하루(20시간)기준 운행빈도 및 배차간격	71
<표 IV-12> 파라미터 추정치	72
<표 IV-13> 서울대입구-고속터미널 간 최적전략($a \in \overline{A}$)	74

<표 IV-14> 서울대입구-도곡 간 최적전략($a \in \overline{A}$)	75
<표 IV-15> 서울대입구-온수 간 최적전략($a \in \overline{A}$)	75
<표 IV-16> 최적전략 알고리즘의 배정통행량과 오차율	76
<표 IV-17> 출발역-도착역 별 경로 및 경로속성	77
<표 IV-18> 경로기반 확률적 알고리즘의 경로선택확률	78
<표 IV-19> 배정통행량과 오차율	78
<표 IV-20> I_i 데이터 구축 예시	82
<표 IV-21> F_i 데이터 구축 예시	82
<표 IV-22> 차내시간 기준 환승통행저항(거리기준)	83
<표 IV-23> 출발역-도착역 간 최단거리 초기값(링크길이)	85
<표 IV-24> 출발역-도착역 간 최단거리 결과값(1)	86
<표 IV-25> 최단경로 이전노드 초기값	86
<표 IV-26> 최단경로 이전노드 결과값	86
<표 IV-27> 출발역-도착역 간 최단거리 결과값(2)	87
<표 IV-28> 서울대입구-온수 간 전체 경로와 효율적 경로	93
<표 IV-29> 서울대입구-도곡 구간의 θ 에 따른 $f(\theta)$ 의 결과	97
<표 IV-30> 환승링크길이(가중치) 설정에 따른 θ^* 의 변화	98
<표 IV-31> 출발지-도착지 별 θ 의 정산 결과	99
<표 IV-32> θ^* 별 링크이용확률 및 배정통행량(고속터미널)	101
<표 IV-33> θ^* 별 구간이용확률 및 배정통행량(도곡)	102
<표 IV-34> θ^* 별 구간이용확률 및 배정통행량(온수)	103
<표 IV-35> 경로선택확률	104
<표 IV-36> 배정통행량과 오차율	105

<표 V-1> 강남-신풍 간 효율적 경로	115
<표 V-2> 신풍-강남 간 효율적 경로	116
<표 V-3> 장승배기-강남(강남-장승배기) 간 효율적 경로	117

그 립 목 차

<그림 I-1> 연구의 흐름도	5
<그림 II-1> 예제 네트워크(1)	11
<그림 II-2> 노선 이용확률	13
<그림 II-3> 예제 네트워크(2)	16
<그림 II-4> 간단히 표현된 예제 네트워크(1)	17
<그림 II-5> 예제 네트워크(3)	19
<그림 III-1> θ 및 최단경로와의 거리차이 별 경로의 선택확률	34
<그림 III-2> 예제 네트워크(4)	36
<그림 III-3> 예제 네트워크(5)	36
<그림 III-4> 예제 네트워크(6)	37
<그림 III-5> 서울대입구-금호 간 전체 네트워크	38
<그림 III-6> 서울대입구-금호 간 효율적 경로	39
<그림 III-7> 예제 네트워크(7)	40
<그림 III-8> 예제 네트워크(8)	41
<그림 III-9> 신대방-서울역 간 전체 경로	43
<그림 III-10> 신대방-서울역 간 효율적 경로	43
<그림 III-11> 예제 네트워크(9)	45
<그림 III-12> 기본적인 네트워크 표현 방식	46
<그림 III-13> 환승통행 반영 네트워크	47
<그림 III-14> Dial 알고리즘을 적용하기 위한 네트워크 표현 방식	48
<그림 III-15> 일반화된 네트워크 표현 방식	49

<그림 III-16> 예제 네트워크(10)	50
<그림 IV-1> 예제 네트워크(11)	57
<그림 IV-2> 링크이용확률(대기시간 가중치=1이상)	58
<그림 IV-3> 최단경로거리와 효율적 경로(대기시간가중치=1)	62
<그림 IV-4> $\theta = 0$ 을 적용하여 구한 링크의 이용확률	63
<그림 IV-5> $\theta = 1$ 을 적용하여 구한 링크의 이용확률	63
<그림 IV-6> $\theta = 10$ 을 적용하여 구한 링크의 이용확률	63
<그림 IV-7> 예제 네트워크(12)	64
<그림 IV-8> 최단경로거리와 효율적 경로(대기시간가중치=1.5)	65
<그림 IV-9> 서울대입구-고속터미널 간 네트워크	68
<그림 IV-10> 서울대입구-고속터미널 간 선택가능 노선(전략)	70
<그림 IV-11> 서울대입구-도곡 간 선택가능 노선(전략)	70
<그림 IV-12> 서울대입구-온수 간 선택가능 노선(전략)	71
<그림 IV-13> 서울대입구-고속터미널 간 네트워크	73
<그림 IV-14> 서울대입구-도곡 간 네트워크	73
<그림 IV-15> 서울대입구-온수 간 네트워크	73
<그림 IV-16> 노드번호 설정(예시)	81
<그림 IV-17> 서울대입구-온수 간 전체 네트워크	92
<그림 IV-18> 서울대입구-온수 간 효율적 경로	92
<그림 IV-19> 정산 알고리즘	96
<그림 V-1> 강남-신풍 간 전체 네트워크와 효율적 경로	115
<그림 V-2> 신풍-강남 간 전체 네트워크와 효율적 경로	116
<그림 V-3> 장승배기-강남(위), 강남-장승배기(아래) 간 효율적 경로	117

I. 서론

1. 연구의 배경 및 목적

대중교통 통행배정 알고리즘은 전통적인 승용차 통행배정 알고리즘의 분류 방법과 마찬가지로 통행자의 인지오차 고려여부에 따라 결정적 통행배정 알고리즘과 확률적 통행배정 알고리즘으로 분류할 수 있다. 또한 분석단위가 링크인지 경로인지에 따라 다시 링크기반 통행배정 알고리즘과 경로기반 통행배정 알고리즘으로 나뉜다.

현재 표준적으로 사용하고 있는 대중교통 통행배정 알고리즘은 링크기반 결정적 통행배정 알고리즘과 경로기반 확률적 통행배정 알고리즘이다. 그런데 두 알고리즘은 각각 통행자의 경로선택에 대한 ‘결정론적’인 가정과 분석단위가 ‘경로’기반이라는 측면에서 결함이 나타난다.

링크기반 결정적 통행배정 알고리즘은 최적링크집합에 전량 통행배정하는 기법으로 최적전략 알고리즘이 대표적이다. 대중교통 이용자는 여러 서비스 노선(common lines)이 통과하는 정차역에서 대기하면서 자신의 총통행시간을 최소화하기 위한 서비스 조합을 선택할 수 있는데, 최적전략 알고리즘은 이러한 통행특성을 설명하기 위하여 최적전략이라는 개념을 도입하였다. 최적전략이란 총예상통행시간을 최소화하는 전략의 집합으로 정의되며 경우에 따라서 다중경로 집합으로 나타난다. 이와 같이 최적전략 알고리즘은 환승통행 등 대중교통 이용자의 통행패턴을 설명할 수 있고 다중경로에 통행배정할 수 있다는 장점 때문에 널리 사용되고 있다. 그러나 최적전략 알고리즘에서 경로의 이용확률은 효용극대화이론을 기반으로 하는 확률적 통행배정 모형의 이용확률과는 다르다. 최적전략 알고리즘에서 경로의 이용확률은 이용자의 인지오차를 고려한 것이 아닌 분석가의 관점에서 정의하는 결정론적인 측면의 이용 ‘비율’이다. 이렇듯 최적전략 알고리즘은 결정적 통행배정 알고리즘으로서 이용자의 관점이 아닌 분석가의 관점에서 대중교통 통행을 설명하기

때문에 아무리 이용자 중심적인 변수를 사용하더라도 이용자의 경로선택에 대한 인지오차를 설명할 수 없다.

경로기반 확률적 통행배정 알고리즘은 통행자들이 같은 경로를 사용하더라도 그 통행시간을 다르게 인지한다는 ‘확률론적’인 가정을 함으로써 앞서 설명한 링크기반 결정적 통행배정 알고리즘의 ‘결정론적’인 가정을 완화한다. 그러나 경로기반 확률적 통행배정 알고리즘은 분석단위를 ‘경로’로 하기 때문에 출발지-도착지 간에 존재하는 모든 경로의 정보가 필요하다. 많은 수의 링크와 노드로 구성된 대규모 네트워크에서는 출발지와 도착지 간에 무수히 많은 경로가 존재하기 때문에 모든 경로를 나열(enumeration)하는 것은 사실상 불가능하다. 경로나열문제를 해결하기 위한 노력으로 전체 경로집합에서 일부 경로만을 추출하는 방법론들이 연구되어 왔지만 경로 선택집합을 설정하는 기준이 임의적이다. 아울러, 경로기반 확률적 통행배정 알고리즘은 효용극대화의 원칙에 따라 가장 높은 효용을 갖는 경로를 선택한다고 가정하기 때문에 경로의 효용함수를 추정해야 한다. 경로선택에 영향을 미치는 요인에는 경로의 서비스 수준을 설명하는 변수(차내시간, 운임, 운행빈도, 환승횟수 등), 통행자의 사회·경제적인 측면을 설명하는 변수(소득, 나이, 성별 등) 등 상당히 많은 변수가 있으며, 이 뿐만 아니라 근래에는 네트워크 토폴로지, 이용자에게 제공되는 네트워크 토폴로지 정보 등도 통행자의 경로선택에 영향을 미치는 유의한 변수로 보고되고 있다. 이러한 점으로부터, 경로선택에 영향을 미치는 속성변수의 설정과 각 속성변수의 파라미터 추정은 상당한 시간과 비용을 수반하며, 연구자 또는 시간적 범위와 공간적 범위의 변화에 따라서 달라질 수 있다는 것을 추론할 수 있다.

링크기반 결정적 통행배정 알고리즘과 경로기반 확률적 통행배정 알고리즘은 앞서 설명한 내재적인 결함에도 불구하고 신중한 검토와 충분한 논의없이 교통수요분석에 적용되고 있다.

이로부터 본 연구는 다음을 주요 연구목적으로 한다. 첫째, 현재 표준적으로 적용되고 있는 대중교통 통행배정 알고리즘인 링크기반 ‘결정

적’ 통행배정 알고리즘과 ‘경로기반’ 확률적 통행배정 알고리즘의 특성과 한계점을 검토하고 시사점을 도출한다. 둘째, ‘링크기반’의 ‘확률적’ 통행배정 알고리즘이 연산 효율적이고 통행자의 경로선택에 대한 무리한 가정을 하지 않는다는 점에서 위의 두 알고리즘보다 합리적인 통행배정 알고리즘임을 제안한다. 특히, 링크기반 확률적 통행배정 알고리즘 중에서 Dial 알고리즘에 대하여 살펴보고, 위의 두 알고리즘과 비교·분석함으로써 링크기반 확률적 알고리즘에 대한 당위성을 논의한다. 셋째, 승용차의 통행배정 알고리즘으로 고안된 Dial 알고리즘을 도시철도 네트워크에 적용 및 분석하기 위하여 고려해야하는 점들과 대중교통 통행배정에서의 적합성에 대해 논의하고자 한다.

2. 연구의 대상 및 범위

본 연구에서는 대중교통 통행배정 알고리즘을 링크기반 결정적 알고리즘, 경로기반 확률적 알고리즘, 링크기반 확률적 알고리즘으로 분류하고, 각 유형 별 대표적인 알고리즘인 최적전략 알고리즘, 로짓기반 알고리즘, 본 연구를 통해 수정 및 보완된 Dial 알고리즘을 각각 평가한다. 성과도 평가는 예제 네트워크와 실제 네트워크에서 이루어 진다.

먼저, 가상적으로 설계된 경로특성과 가중치를 적용한 예제 네트워크를 기반으로 각 알고리즘을 비교, 분석한다. 예제 네트워크는 Spiess and Florian(1989)가 최적전략 알고리즘을 설명하기 위해 사용한 네트워크를 차용한다. 예제 네트워크를 기반으로 각 알고리즘의 특성을 살펴본 후에는 이를 토대로 실제 네트워크에 적용해본다. 실제 네트워크는 수도권 도시철도 네트워크 중 일부 출발지-도착지 쌍으로 선정하며, 알고리즘의 성과도 평가는 실제 통행량 자료와 각 알고리즘의 배정통행량 결과를 비교하며 수행한다.

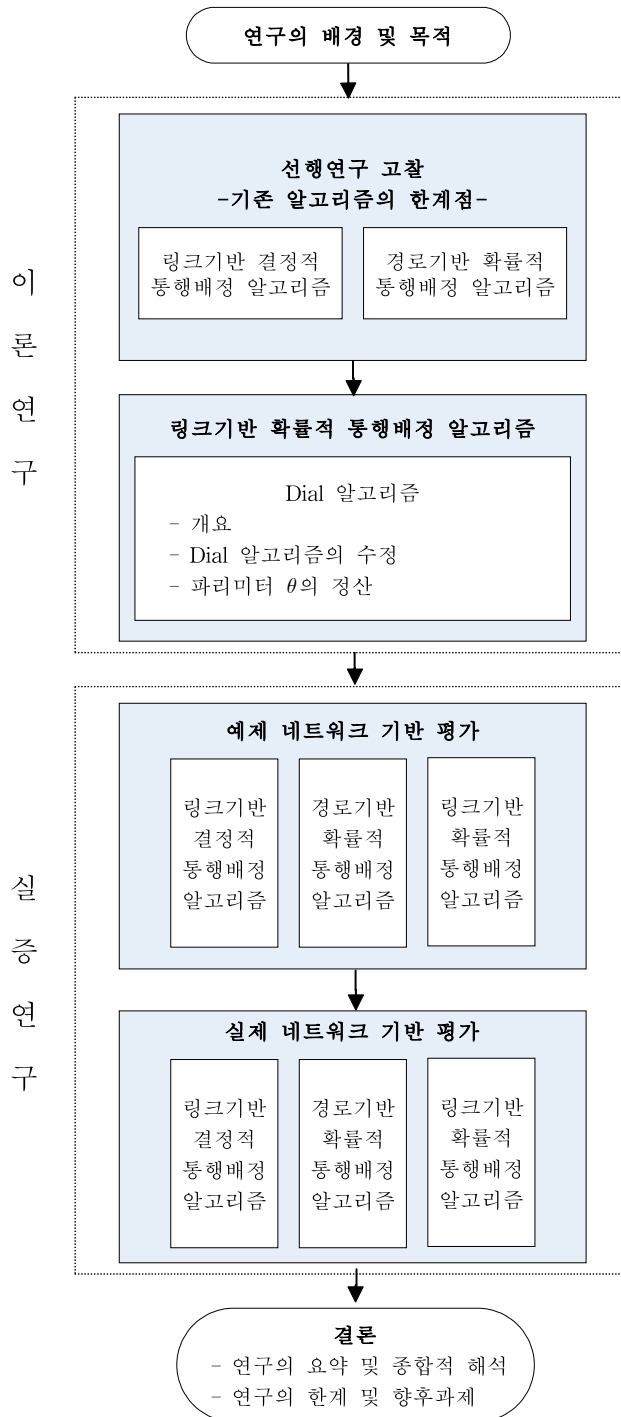
3. 연구의 구성 및 방법

본 연구는 II장과 III장의 이론 연구, IV장의 실증 연구, V장의 종합적 해석의 세 부분으로 구성된다. 먼저 II장의 선행연구 고찰을 통해 현재 대중교통 통행배정 알고리즘으로 주로 사용하고 있는 링크기반 결정적 알고리즘과 경로기반 확률적 알고리즘의 특성과 한계점을 검토한다. III장에서는 이 두 알고리즘의 극복 대안으로 링크기반 확률적 알고리즘을 제안한다. 특히, 링크기반 확률적 통행배정 알고리즘으로 대표적인 Dial 알고리즘에 대해서 살펴본다. 또한, 승용차 통행배정을 목적으로 고안된 Dial 알고리즘을 대중교통 통행배정에 적용할 때 고려해야하는 점, 특히 Dial 알고리즘에 환승통행을 합리적으로 반영하기 위한 네트워크 표현방식과 특정 출발지-도착지 간 경로분석 방법 및 알고리즘의 정산 방법 등을 논의한다.

IV장의 실증 연구는 예제 네트워크와 실제 네트워크에서 각 알고리즘의 성과도를 평가한다. 우선 가상적인 예제 네트워크에 링크기반 결정적 알고리즘, 경로기반 확률적 알고리즘, 링크기반 확률적 알고리즘을 적용하여 결과를 비교·분석한다. 앞서 II장과 III장의 이론 연구에서 살펴본 각 알고리즘의 특성 및 한계점을 실증적으로 점검하고 링크기반 확률적 알고리즘의 적합성을 논의한다. 다음으로 실제 수도권 도시철도 네트워크에 각 알고리즘을 적용하는데, 각 알고리즘 별로 적용상의 특징 및 유의점을 살펴보고, 알고리즘이 대중교통 통행특성을 얼마나 묘사할 수 있는지 실제 통행량 자료와 비교하여 평가한다.

마지막으로 V장의 종합적 해석을 통해 본 연구의 결과를 요약 및 정리하고 대중교통 통행배정 알고리즘으로서 Dial 알고리즘의 적합성에 대한 논의를 거친 뒤 향후 연구과제를 설정한다.

위에서 기술한 일련의 연구 절차는 <그림 I-1>과 같다.



<그림 I-1> 연구의 흐름도

II. 선행연구 고찰

1. 대중교통 통행배정 알고리즘의 개요

통행배정 알고리즘은 출발지-도착지(O-D)간의 통행량을 네트워크에 부하(loading)하는 기법이다. 통행배정 알고리즘의 분류 기준은 다양하나 주로 혼잡효과 고려여부와 통행자의 인지오차 고려여부로 구분한다.

먼저, 혼잡효과 고려여부에 따라 네트워크 균형(equilibrium) 기법과 네트워크 부하(network loading) 기법으로 분류할 수 있다. 네트워크 균형(equilibrium) 기법은 링크의 혼잡효과를 고려하여 링크의 통행비용이 통행량의 함수로 정의되고, 반복적 기법을 통해 네트워크의 균형(equilibrium) 상태를 찾는다. 반면에 네트워크 부하(network loading) 기법은 혼잡효과를 고려하지 않아 링크의 통행비용이 통행량과 관계없이 일정하며, 균형 기법과 구분하기 위하여 네트워크 부하(network loading) 기법이라고 한다(Daganzo and Sheffi, 1977).

또한, 통행자의 인지오차 고려여부에 따라 결정적(deterministic) 알고리즘과 확률적(stochastic) 알고리즘으로 분류할 수 있다. 결정적 알고리즘은 모든 통행자가 네트워크의 교통상황을 정확히 알고 있으며, 실제 통행비용과 통행자가 인지하는 통행비용 간의 오차가 없다고 가정한다. 즉, 모든 통행자가 경로를 선택할 때 통행비용을 기준으로 완벽한 의사결정을 한다고 가정한다. 이와 달리, 확률적 알고리즘은 통행자들이 경로 선택 시에 단일 기준으로 완벽한 의사결정을 하는 것이 아니라 같은 경로를 통행하더라도 그 통행비용을 다르게 인지한다고 가정한다.

통행배정 알고리즘은 위의 두 가지 분류 기준에 따라 <표 II-1>과 같이 구분할 수 있다. 먼저 혼잡효과를 고려하지 않고, 모든 통행자가 경로의 통행시간을 정확하게 인지하여 최단경로로만 통행하는 것을 가정하는 전량배정기법이 있다. 전량배정기법은 계산적으로 효율적이지만 통행

시간의 작은 변화에도 통행배정 결과가 민감하게 달라질 수 있고, 통행자가 동일하고 정확하게 통행비용을 인지한다는 가정이 비현실적이다. 두 번째로 혼잡효과는 고려하지 않지만 통행자의 인지오차를 반영하는 확률적 네트워크 부하(Stochastic Network Loading: SNL) 알고리즘이 있다(Daganzo and Sheffi, 1977). SNL 알고리즘은 주로 로짓모형 또는 프로빗 모형을 기반으로 한다. 세 번째로 통행자가 정확하게 통행비용을 인지하고 최단경로로 통행한다고 가정하는 점에서 전량배정기법과 동일하나, 혼잡효과를 고려하는 이용자균형 통행배정 알고리즘이 있다. 이용자균형 통행배정 알고리즘은 이용자가 경로를 변경하더라도 그들의 통행시간을 줄일 수 없는 ‘균형’상태의 결과를 도출할 때까지 반복적으로 계산한다. 마지막으로 확률적 이용자균형 통행배정 알고리즘은 이용자균형 통행배정 알고리즘에서 통행자의 인지오차를 반영한 알고리즘이다. SNL 알고리즘에서 통행시간을 통행량의 함수로 구성한 후 반복적 기법을 통해 혼잡효과를 반영한 알고리즘이라고도 해석할 수 있다.

<표 II-1> 통행배정 알고리즘의 분류

구분		통행시간 인지오차 고려여부	
		결정적 통행배정	확률적 통행배정
혼잡효과 고려여부	비고려	전량배정기법 (all-or-nothing)	확률적 네트워크 부하 (stochastic network loading)
	고려	이용자균형 통행배정 (user equilibrium)	확률적 이용자균형 통행배정 (stochastic user equilibrium)

위의 분류 방법은 승용차와 대중교통의 통행배정 알고리즘에 모두 적용할 수 있다. 승용차와 대중교통 통행배정 알고리즘에서 혼잡효과 및 통행자의 인지오차를 반영하는 원리가 크게 차이나지 않기 때문이다. 다만, 대중교통 통행배정 알고리즘은 승용차 이용자와는 다른 대중교통 이용자의 통행특성과 승용차의 공로 시스템과는 다른 대중교통 시스템을

반영하기 위해 두 가지 측면에서 차이가 나타난다.

먼저, 대중교통 이용자의 최단경로(최적경로)는 다중경로로 나타날 수 있다. 왜냐하면, 대중교통 이용자는 여러 서비스 노선(common lines)이 통과하는 정차역에서 대기하며 자신의 총통행시간을 최소화하기 위한 서비스 노선을 선택할 수 있기 때문이다. 이러한 통행특성 때문에 대중교통 이용자의 최적경로는 승용차 이용자의 최단경로와는 달리 각각의 기대확률을 갖는 다수의 경로 또는 노선의 집합일 수 있다. 대표적인 예는 Spiess and Florian(1989)의 최적전략 알고리즘으로, 출발지-도착지 간 ‘기대(expected)’ 총통행시간을 최소화하는 최적경로를 승용차 이용자의 최단경로와 구분하기 위하여 최적전략(optimal strategies)이라는 개념으로 제시한 바 있다. 이처럼 대중교통 통행배정 알고리즘에서는 최적경로를 단일경로가 아닌 다중경로로 정의할 수 있다는 점 때문에 대중교통 통행배정 알고리즘을 구분할 때 혼돈이 발생하기도 한다. 예컨대, 임강원·임용택(2003)은 대중교통 통행배정 알고리즘을 크게 전량통행배정기법, 최적전략기반 통행배정 기법, 확률적 통행배정 기법으로 구분하고 있다. 최적전략 알고리즘 역시 최적전략에 전량통행배정하는 결정적 알고리즘이라는 점을 감안한다면 전량통행배정기법에 포함시켜 구분(Liu, Y. et al., 2010)하는 것이 보다 바람직할 것이다.

둘째, 대중교통 통행배정 알고리즘에서 고려하는 혼잡효과는 승용차 통행배정 알고리즘에서의 혼잡효과와 다소 차이가 있다. 승용차 통행배정 알고리즘에서 다루는 주된 혼잡효과는 링크의 통행시간 변화를 유발하는 ‘링크 상에서의 지체’이다. 반면에, 궤도교통이나 BRT와 같이 타수단의 통행량에 영향을 받지 않고, 규칙적인 배차간격으로 고정된 노선을 운행하는 대중교통 시스템의 경우 혼잡으로 인한 ‘링크 상에서의 지체’가 발생하지 않는다. 대신 대중교통 통행배정 알고리즘은 승·하차 시 혼잡과 이로 인한 출발지체, 차량 내 혼잡과 이로 인한 서비스 수준 저하 등을 혼잡효과로 고려할 수 있다. 한편, 이러한 혼잡효과는 알고리즘의 통행량 부하방식이 대중교통 이용자의 경로선택 특성에 부합하는지에

대한 논의가 선행된 후 반복적 기법을 통해 반영이 가능하다.

지금까지 통행배정 알고리즘의 분류와 대중교통 통행배정의 두 가지 특성에 대하여 살펴보았다. 요약하면, 첫째, 통행배정 알고리즘은 혼잡효과 고려여부와 통행자의 통행시간 인지오차 고려여부에 따라 분류할 수 있다. 둘째, 대중교통 통행배정 알고리즘은 승용차 통행배정 알고리즘과는 달리 최적경로가 다중경로로 나타날 수 있다는 점에서 차이가 있으나 알고리즘의 분류 방법은 동일하다. 셋째, 대중교통 중 궤도교통은 혼잡으로 인한 링크 상의 지체가 발생하지 않으며, 기타 혼잡효과와의 반영은 알고리즘의 부하방식에 대한 논의가 선행된 후 반영가능하다.

따라서 본 연구는 연구 범위를 <표 II-2>와 같이 혼잡효과를 고려하지 않은 결정적, 확률적 통행배정 알고리즘으로 한정한다. 또한 알고리즘을 다시 분석단위(링크 또는 경로)에 따라 링크기반 알고리즘과 경로기반 알고리즘으로 분류하고, 각 알고리즘의 통행량 부하방식이 수도권 도시철도 이용자의 통행특성에 부합하는지에 대하여 중점적으로 논의하고자 한다.

<표 II-2> 연구의 범위

구분	결정적 통행배정	확률적 통행배정
종류	<ul style="list-style-type: none"> - 링크(전략)기반 전량배정기법 - 경로기반 전량배정기법 	<ul style="list-style-type: none"> - 링크기반 확률적 네트워크 부하 - 경로기반 확률적 네트워크 부하

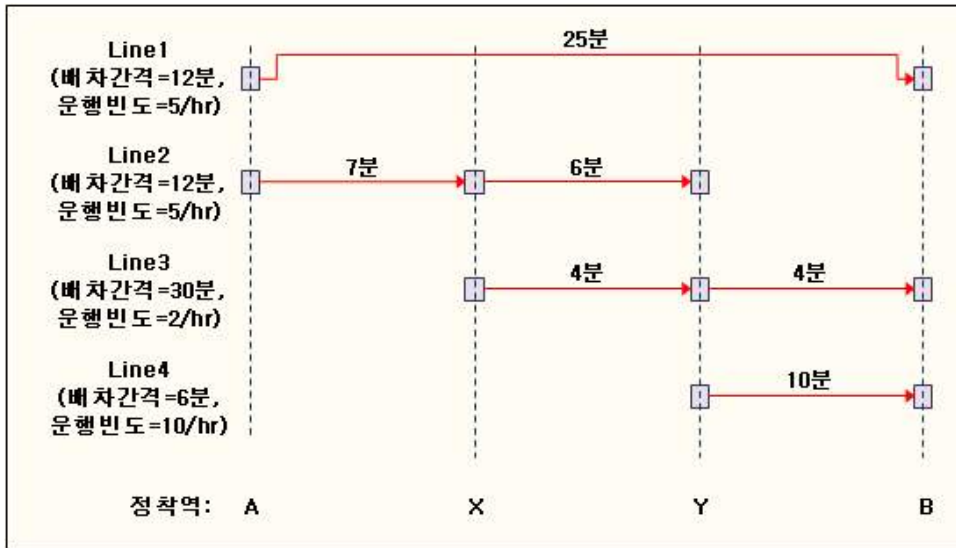
다음의 2절과 3절에서는 현재 대중교통 통행배정에서 주로 사용하고 있는 링크기반 결정적 통행배정 알고리즘과 경로기반 확률적 통행배정 알고리즘에 대하여 살펴본다.

2. 링크기반 결정적 알고리즘

링크기반 결정적 알고리즘은 최적링크집합에 전량을 배정하는 기법으로서 Spiess and Florian(1989)의 최적전략 알고리즘과 교통수요분석 툴인 TransCad에 탑재된 Pathfinder가 대표적이다.

Spiess and Florian(1989)의 최적전략 알고리즘은 승용차 이용자와는 다른 대중교통 이용자의 통행특성을 설명하기 위해서 고안된 알고리즘이다. 승용차 이용자는 자신의 통행시간을 최소화할 수 있는 단일 경로를 선택하는 반면에, 대중교통 이용자는 여러 서비스 노선(common lines)이 통과하는 정차역에서 대기하면서 자신의 총통행시간을 최소화하기 위한 서비스 조합을 선택할 수 있다. 따라서 지정된 단일 경로를 선택하여 통행하는 것보다 더 복잡한 문제이다. 최적전략 알고리즘은 이러한 대중교통 이용자의 통행특성을 알고리즘 상에 나타내기 위하여 전략(strategy)과 최적전략(optimal strategy)이라는 개념을 도입하였다. 전략은 출발지로부터 목적지까지 통행하는 동안 이용자가 선택할 수 있는 승·하차, 차내 통행, 환승, 대기 등의 모든 통행요소를 의미하고, 최적전략은 대기시간을 포함한 총예상통행시간을 최소화하는 전략의 집합으로 정의한다. 최적전략 알고리즘에서 모든 통행은 최적전략에 포함되는 전략에 한하여 배정되고 최적전략이 아닌 전략의 이용확률은 0이다.

전략의 개념은 Spiess and Florian(1989)의 예제 네트워크(1)을 통해 보다 쉽게 이해할 수 있다. <그림 II-1>의 예제 네트워크(1)에는 네 개의 정차역 A, X, Y, B와 정차역 간 차내시간, 각 정차역을 통과하는 서비스 노선과 배차간격 정보가 제시되어 있다. 정차역 A에서는 'Line1', 'Line2' 서비스 노선을, 정차역 X에서는 'Line2', 'Line3' 서비스 노선을, 정차역 Y에서는 'Line3', 'Line4' 서비스 노선을 이용할 수 있다. 이용자



Spiess and Florian(1989)

<그림 II-1> 예제 네트워크(1)

가 서비스를 이용하기 위한 예상대기시간은 서비스 노선의 운행빈도 합수로 나타나며, 이용가능한 서비스가 2개 이상일 경우 각 서비스 노선의 운행빈도를 선형 결합한 값의 역수, 즉 결합 예상대기시간을 대기시간으로 사용한다.

$$W(\overline{A_i^+}) = \frac{\alpha}{\sum_{a \in A_i^+} f_a}, \quad \alpha > 0$$

여기서, a 는 링크집합 A 에 속하는 링크(i, j), f_a 는 링크 a 의 서비스 운행빈도, $\overline{A_i^+}$ 는 노드 i 에서의 최적전략, $W(\overline{A_i^+})$ 는 링크 a 의 결합대기시간이다. α 는 배차간격에서 대기시간의 비율을 의미하는 동시에 대기시간에 대한 차내시간 가중치를 의미한다. 예를 들어, 배차간격의 절반이 대기시간이고, 차내시간 기준 대기시간 가중치가 2라고 가정하면, α 는 1.0 ($=0.5 \times 2$)가 된다. 본 예제에서는 α 를 0.5로 가정한다.

A, X, Y 정차역에서 출발하여 B 정차역까지 도달하는데 총

75(5×5×3)개의 전략이 있으며 각 정차역에서 선택가능한 전략, 예상대기 시간, 노선의 이용확률은 <표 II-3>과 같이 나타난다. 최적전략 알고리즘에서 노선의 이용확률은 노선의 운행빈도를 결합운행빈도로 나눈 값이다.

$$P_a(\overline{A_i^+}) = \frac{f_a}{\sum_{a' \in \overline{A_i^+}} f_{a'}}, \quad a \in \overline{A_i^+}$$

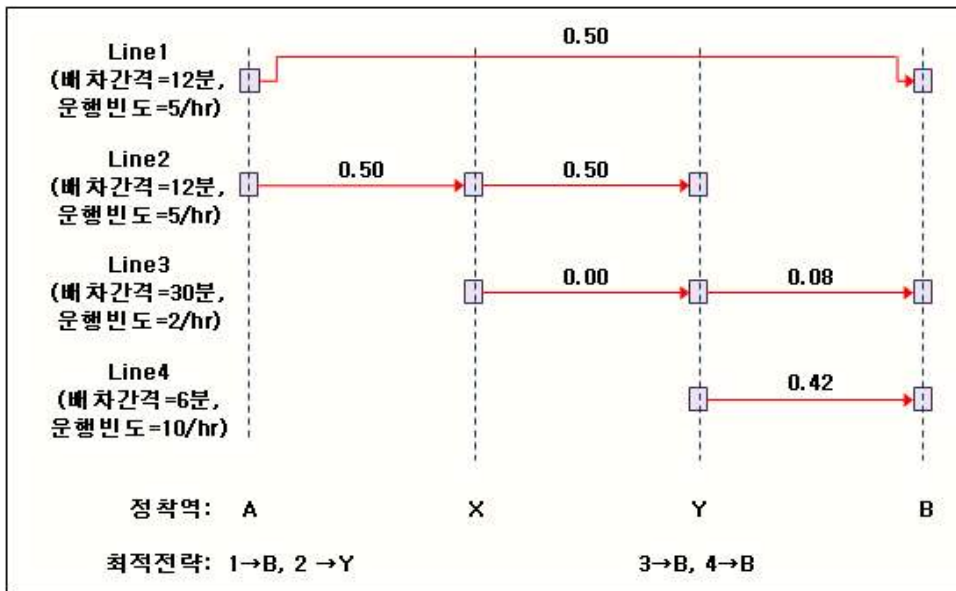
<표 II-3> 각 출발역에서 도착역 B까지의 전략, 대기시간 및 이용확률

출발역	전략(노선→하차역)	대기시간	노선 이용 확률			
			Line1	Line2	Line3	Line4
A	1 → B	6	1.00	-	-	-
A	2 → X	6	-	1.00	-	-
A	2 → Y	6	-	1.00	-	-
A	1→B, 2→X	3	0.5	0.5	-	-
A	1→B, 2→Y	3	0.5	0.5	-	-
X	2 → Y	6	-	1.00	-	-
X	3 → Y	15	-	-	1.00	-
X	3 → B	15	-	-	1.00	-
X	2 → Y, 3 → Y	4.3	-	0.71	0.29	-
X	2 → Y, 3 → B	4.3	-	0.71	0.29	-
Y	3 → B	15	-	-	1.00	-
Y	4 → B	3	-	-	-	1.00
Y	3 → B, 4 → B	2.5	-	-	0.17	0.83

Spiess and Florian(1989)

음영으로 표시된 전략은 출발역 A에서 도착역 B까지의 총예상통행 시간을 최소로 하는 최적전략을 의미한다. 즉, 최적전략은 출발역 A에서 Line1 또는 Line2 중 먼저 도착하는 노선에 승차하여 각각 B 또는 Y 정

차역에 도착하고, Y 정차역에 도착한 경우 Line3 또는 Line4 중에 먼저 오는 노선으로 환승하는 경로집합을 의미한다. 최적전략의 총예상통행시간은 $27.75\text{분} = (7+6+3) \times 0.5 + (25+3) \times 0.5 + (4+2.5) \times 0.17 \times 0.5 + (10+2.5) \times 0.83 \times 0.5$)으로서 전략 조합들 중 최소 통행시간이며, 승용차 통행배정 알고리즘에서 최단경로 통행시간인 31분(=25+6)보다 작은 값이다. 최적전략에 포함된 각 전략이 이용될 확률은 <그림 II-2>와 같다.



Spiess and Florian(1989)

<그림 II-2> 노선 이용확률

위의 예제를 통해 알 수 있듯이, 최적전략 알고리즘은 대기시간을 포함한 총예상통행시간을 최소로 하는 최적전략을 찾고, 이용확률을 구한다. 이러한 일련의 과정은 후방진행단계와 전방진행단계를 통해 진행된다. 먼저, 첫 번째 단계는 후방진행단계로서, 도착지에서 시작하여 모든 출발지까지의 총예상통행시간을 구하고 최적전략을 찾는다. 두 번째 단계인 전방진행단계에서는 출발지에서 시작하여 도착지에 도달할 때까지 최적전략에 속하는 각 전략의 이용확률을 구한다.

0. 초기화 단계: 도착지(d)를 제외한 모든 노드 i 에 대하여 예상통행 시간 u_i 와 노드 i 의 결합 운행빈도 f_i 를 초기화한다.

$$u_i = \begin{cases} 0 & \text{if } i = d \\ \infty & \text{otherwise} \end{cases}$$

$$f_i = 0$$

1. 후방진행 단계: 한 번도 탐색하지 않은 링크 중에서 $u_j + c_a$ 값이 최소인, 즉 도착지로부터 가장 가까운 링크 $a(i, j)$ 부터 선택하여 $u_i \geq u_j + c_a$ 이면 최적전략에 포함하고 u_i 와 f_i 를 갱신한다. 모든 링크를 탐색한 후, 두 번째 단계로 넘어간다. 여기서 c_a 는 링크 $a(i, j)$ 의 통행시간, $u_j + c_a$ 는 노드 i 에서 대기시간을 제외한 도착지까지의 통행시간이고 u_i 는 대기시간을 포함한 도착지까지의 통행시간이다. .

$$\overline{A_i^+} = \overline{A_i^+} + \{a\},$$

$$u_i = \begin{cases} \{f_i u_i + f_a(u_j + c_a)\} / (f_i + f_a) & \text{if } f_a < \infty \\ u_j + c_a & f_a = \infty \end{cases},$$

$$f_i = \begin{cases} f_i + f_a & \text{if } f_a < \infty \\ \infty & f_a = \infty \end{cases}.$$

2. 전방진행 단계: $u_j + c_a$ 값이 최대인, 즉 출발지로부터 가까운 링크 부터 이용확률을 구하고 배정 통행량을 구한다.

$$P_a = \begin{cases} f_a / f_i & \text{if } a \in \overline{A_i^+} \\ 0 & \text{otherwise} \end{cases},$$

$$v_a = v_a + P_a V_i,$$

$$V_j = V_j + P_a V_i.$$

Spiess and Florian(1989)는 대기시간을 배차간격의 절반으로 사용하
 되, 표현을 간단히 하기위하여 α 를 1로 가정하고 대신 각 운행빈도 f 를
 실제 값의 2배로 적용하였다. u_i 를 α 값을 사용하여 다시 나타내면 식
 (1)과 같다.

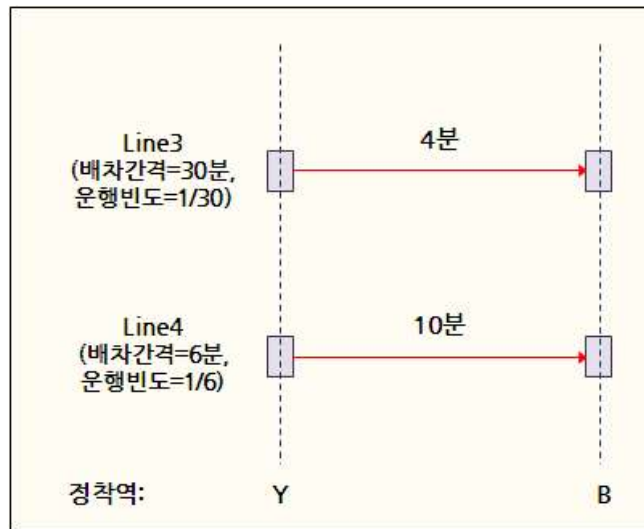
$$\begin{aligned}
 u_i &= \frac{f_i}{(f_i + f_a)} u_i + \frac{f_a}{(f_i + f_a)} (u_j + c_a) \\
 &= \frac{f_i}{(f_i + f_a)} \left(\frac{\alpha}{f_i} + u'_i \right) + \frac{f_a}{(f_i + f_a)} (u_j + c_a) \\
 &= \frac{\alpha}{(f_i + f_a)} + \frac{f_i}{(f_i + f_a)} u'_i + \frac{f_a}{(f_i + f_a)} (u_j + c_a).
 \end{aligned} \tag{1}$$

여기서 u'_i 는 노드 i 의 대기시간을 포함하지 않는 노드 i 에서 도착지
 까지의 예상통행시간이다. 식 (1)의 우항에서 첫 번째 항은 노드 i 에서의
 결합대기시간을, 두 번째와 세 번째 항의 합은 노드 i 에서 도착지까지
 최적전략의 예상통행시간이며 노드 i 의 대기시간은 포함하지 않는다.

최적전략 알고리즘과 함께 대표적인 링크기반 결정적 대중교통 통행
 배정 알고리즘으로 Pathfinder가 있다. Pathfinder는 TransCad에서 제공
 하는 핵심적인 통행배정 알고리즘 중 하나로서 최적전략 알고리즘의 일
 반화된 방법이다. 최적전략 알고리즘과 동일한 로직으로 최적전략 집합
 을 구축하고 노선의 운행빈도 비율에 비례하여 노선의 이용확률을 구한
 다. 두 알고리즘의 차이점은 최적전략 알고리즘이 통행시간을 척도로 최
 적경로를 탐색하는 반면, Pathfinder는 통행비용까지 고려한 일반화 비용
 을 척도로 사용한다. 또한, Pathfinder는 환승횟수 제한 등의 추가적인
 변수를 고려할 수 있다.

최적전략 알고리즘과 Pathfinder는 여러 서비스 노선이 통과하는 정
 차역에서 대중교통 이용자의 노선선택 및 환승통행 등을 설명할 수 있
 고, 다중경로에 통행배정할 수 있다는 장점 때문에 널리 사용되고 있다.

그러나 최적전략 알고리즘은 기본적으로 한 지점에서 먼저 도착하는 서비스 노선에 탑승하여 통행한다는 것을 가정하기 때문에 승·하차 및 환승이 한 정류장에서 가능한 버스이용자의 통행배정에 적합하고, 플랫폼이 공간적으로 떨어져 있어 한 지점에서 환승할 수 없는 지하철 이용자의 통행배정에는 부적합하다. 이러한 모형의 가정은 차치하더라도, 최적전략을 구축할 때 대기시간을 제대로 고려하지 못하고, 링크의 이용확률이 운행빈도만의 함수로 이루어져 다른 서비스수준은 반영하지 못한다는 단점이 있다. 이 점은 <그림 II-3>의 예제 네트워크(2)에서 노선이용확률을 산정해봄으로써 확인할 수 있다. 예제 네트워크(2)는 예제 네트워크(1)의 일부 구간이며 가상적으로 Line3과 Line4에 각각 차내시간 기준 대기시간 가중치인 α , β 를 적용한다.

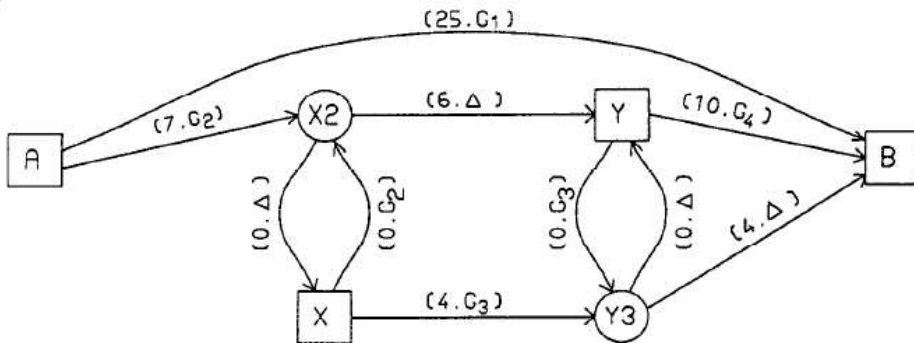


<그림 II-3> 예제 네트워크(2)

우선 후방진행단계에서 $u_j + c_a$ 값이 4분으로 최소값을 갖는 Line3의 링크를 선택하고 u_Y 를 $u_Y = \infty$ 에서 $u_Y = 4 + (1/30)^{-1} \times 0.5 \times \alpha = 4 + 15\alpha$ 로 갱신한다. 다음으로 $u_j + c_a$ 값이 10분인 Line4의 링크가 선택되고 기존의 u_Y 값과 선택된 링크의 $u_j + c_a$ 를 비교하여 u_Y 가 크거나 같을 경우 최적

전략이 된다. 즉, $4 + 15\alpha \geq 10$ 일 경우 Line4는 최적전략에 포함된다. α 가 0.4 이상이면 Line4의 대기시간 가중치 β 값과 관계없이 Line4는 최적전략에 포함되고, Line3과 Line4의 이용확률은 두 노선의 운행빈도 비율인 1:5로 일정하여 비합리적이다.

<그림 II-1>의 예제 네트워크(1)에서 대기시간 가중치를 α 로 동일하게 적용하더라도, α 값이 1보다 클 경우 그 값과 관계없이 <그림 II-2>와 동일한 최적전략이 구축되고 운행빈도는 변화가 없으므로 노선의 이용확률 역시 똑같이 나타난다. 이 결과는 대기시간 비율이 큰 Line2, Line3, Line4로 구성된 경로와 대기시간 비율이 작은 Line1의 경로의 총 통행시간 차이가 대기시간 가중치에 따라 달라지더라도 그 확률은 운행빈도 비율에 따라 고정되게 나타나는 것을 의미한다. 이는 상식적으로 비합리적인 통행배정 결과이다. <표 II-4>와 <표 II-5>는 <그림 II-1>의 예제 네트워크를 간단히 나타낸 <그림 II-4>에 대기시간 가중치로 2와 100을 적용했을 때의 최적전략(음영부분)을 나타내며, 대기시간 가중치가 1인 <그림 II-2>의 최적전략과 동일하게 나타난다.



Spiess and Florian(1989)

<그림 II-4> 간단히 표현한 예제 네트워크(1)

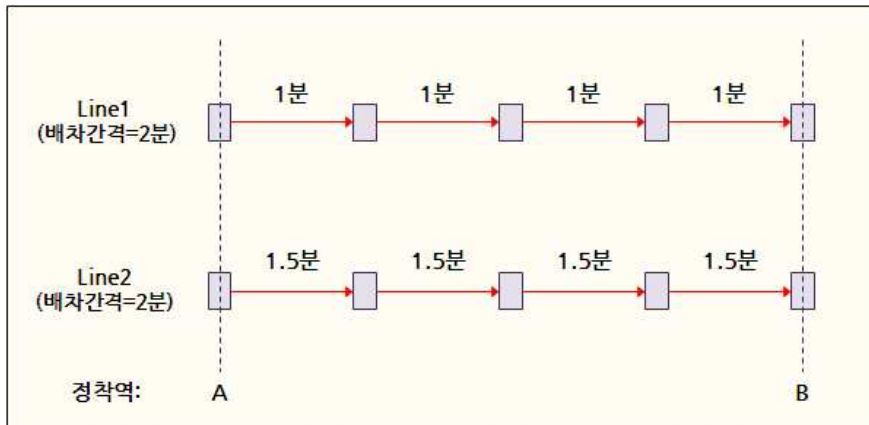
<표 II-4> 대기시간 가중치가 2일 때의 최적전략

#	Node labels (u_i, f_i)						Link with Min $u_j + c_a$			
	A	X2	X	Y3	Y	B	$a(i, j)$	f_a	$u_j + c_a$	$a \in \overline{A}$
1	$\infty, 0$	$\infty, 0$	$\infty, 0$	$\infty, 0$	$\infty, 0$	$0, 0$	Y3, B	∞	4	y
2	"	"	"	$4, \infty$	"	"	Y, Y3	1/15	4	y
3	"	"	"	"	$34, 1/15$	"	X, Y3	1/15	8	y
4	"	"	$38, 1/15$	"	"	"	Y, B	1/3	10	y
5	"	"	"	"	$14, 2/5$	"	Y3, Y	∞	14	n
6	"	"	"	"	"	"	X2, Y	∞	20	y
7	"	$20, \infty$	"	"	"	"	X, X2	1/6	20	y
8	"	"	$25, 7/30$	"	"	"	X2, X	∞	25	n
9	"	"	"	"	"	"	A, B	1/6	25	y
10	$37, 1/6$	"	"	"	"	"	A, X2	1/6	27	y
	$32, 1/3$	$20, \infty$	$25, 7/3$	$4, \infty$	$14, 2/5$	$0, 0$	-	-	-	-

<표 II-5> 대기시간 가중치가 100일 때의 최적전략

#	Node labels (u_i, f_i)						Link with Min $u_j + c_a$			
	A	X2	X	Y3	Y	B	$a(i, j)$	f_a	$u_j + c_a$	$a \in \overline{A}$
1	$\infty, 0$	$\infty, 0$	$\infty, 0$	$\infty, 0$	$\infty, 0$	$0, 0$	Y3, B	∞	4	y
2	"	"	"	$4, \infty$	"	"	Y, Y3	1/15	4	y
3	"	"	"	"	$1504, 1/15$	"	X, Y3	1/15	8	y
4	"	"	$1508, 1/15$	"	"	"	Y, B	1/3	10	y
5	"	"	"	"	$259, 2/5$	"	A, B	1/6	25	y
6	$625, 1/6$	"	"	"	"	"	Y3, Y	∞	259	n
7	"	"	"	"	"	"	X2, Y	∞	265	y
8	"	$265, \infty$	"	"	"	"	X, X2	1/6	265	y
9	"	"	"	"	"	"	A, X2	1/6	272	y
10	$448.5, 1/3$	"	"	"	"	"	X2, X	∞	1508	n
	$448.5, 1/3$	$265, \infty$	$1508, 1/15$	$4, \infty$	$259, 2/5$	$0, 0$	-	-	-	-

또한 <그림 II-1>의 예제 네트워크(1)과는 달리, 실제 지하철 노선들의 배차간격은 거의 비슷하여 차이가 작고 경로통행시간에 비해 작은 비율을 차지한다. 이러한 실제 지하철 네트워크에 최적전략 알고리즘을 적용할 경우 단일 최단경로에 전량 통행배정될 가능성이 높다. 예를 들어, <그림 II-5>의 예제 네트워크(3)과 같이 두 기종점 A, B간 총통행시간이 각각 4분과 6분으로 2분 차이가 나는 Line1과 Line2의 두 경로가 있을 때, 최적전략 알고리즘은 최적전략의 정의에 의해서 단일 최적경로인 Line1에 전량 배정하여 Line2를 이용하는 통행자를 사전적으로 배제한다.



<그림 II-5> 예제 네트워크(3)

이와 같은 문제점은 모든 이용자가 네트워크의 교통상황을 정확히 알고 있으며, 실제 통행비용과 인지 통행비용의 오차가 없이 완벽한 의사결정을 한다는 결정론적 가정에 기인한다. 최적전략 알고리즘과 Pathfinder에서 최적경로가 경우에 따라서 다중경로 집합으로 정의되고 각 경로는 이용확률을 갖지만, 이는 효용극대화이론을 기반으로 하는 확률적 통행배정 모형의 이용확률과는 다르다. 확률적 통행배정에서 경로의 이용확률은 이용자의 인지오차를 고려하는 경로의 선택확률인 반면에, 대중교통 전량배정기법의 이용확률은 이용자 별로 달라질 수 있는 확률이 아닌, 분석가의 관점에서 정의하는 결정론적인 측면의 이용 ‘비

을'이다. 이렇듯 최적전략 알고리즘을 비롯한 전량배정기법은 결정론적 가정을 전제함으로써 경로선택에 대한 이용자의 인지오차를 반영하지 못하고, 아무리 이용자 중심적인 변수를 사용하더라도 분석가의 관점에서 대중교통 이용자의 통행특성을 설명하는 것이다. 또한 통행시간이 아닌 다른 기준에 의해 경로를 선택하는 비결정론적인 통행을 예측하는 것에 한계가 있다. 최적전략 알고리즘과 Pathfinder도 대중교통 이용자의 통행 행태를 설명하기 위하여 승차시간, 승차시간 가중치, 대기시간 가중치, 차두간격에 대한 대기시간 비율, 접근시간 가중치 등 다수의 통행요소 시간과 가중치 변수를 사용하고는 있지만 전량통행배정방식의 기본적인 가정때문에 이용자의 확률적 경로선택을 고려할 수 없다. 오히려 이용자가 노선을 선택할 확률을 노선의 운행빈도 비율로 계산함으로써 다른 서비스 특성보다 운행빈도에 더 큰 가중치를 두는 분석가 중심적인 알고리즘이다. 또한, 다수의 파라미터는 분석의 시·공간적 범위가 다를 경우 범용적 적용이 어려워 충분한 선행연구가 수반되어야 한다는 어려움이 있다.

3. 경로기반 확률적 알고리즘

경로기반 확률적 알고리즘은 분석단위가 ‘경로기반’이고 통행자의 경로선택에 대한 가정이 ‘확률론적’이라는 두 가지 특성을 갖는다.

먼저 ‘확률론적’ 특성은 앞서 살펴본 결정적 통행배정 알고리즘이 ‘모든 통행자가 네트워크의 교통상황을 정확히 알고 있으며, 동일한 기준으로 경로를 선택한다’고 가정하는 것을 완화하기 위하여 제안되었다. Nielsen(2000)은 통행자가 확률적 경로선택을 하는 네 가지 이유를 첫째, 통행자는 네트워크에 대한 완벽한 지식이 없으므로 각자 인지한 효용(perceived utilities)에 따라 합리적으로 경로를 선택하고, 둘째, 대안 경로들의 통행시간은 매번 변할 수 있으며, 셋째, 각각의 통행자는 경로선택에 대한 다른 선호를 가지며, 마지막으로 기타 변동사항에 의해 다른 경로가 선택될 수 있기 때문이라고 정리하였다. 요컨대 확률적 통행배정 알고리즘은 통행자들이 경로 선택 시 단일 기준으로 완벽한 의사결정을 하는 것이 아니고, 같은 경로를 통행하더라도 그 통행비용을 다르게 인지한다고 가정함으로써 결정적 통행배정 알고리즘에 비해 이용자의 통행행태를 합리적으로 설명하고자 한다.

또한 확률적 통행배정 알고리즘에서 통행자는 각 경로에 대하여 각기 다른 효용을 갖고, 효용극대화의 원칙에 따라 가장 높은 효용을 갖는 경로를 선택한다고 가정한다. 일반적으로 효용은 관측이 가능한 고정효용과 관측이 불가능한 확률효용의 두 가지 요소로 구분된다.

$$U_n^P = V_n^P + \epsilon_n^P$$

U_n^P : 개별 통행자 n 의 경로 P 의 총 효용

V_n^P : 개별 통행자 n 의 경로 P 의 고정효용

ϵ_n^P : 개별 통행자 n 의 경로 P 의 확률효용

여기서 확률효용에 대한 분포를 검블(Gumbel)분포로 가정하면 로짓 기반(logit-based) 알고리즘이 유도되고, 정규분포를 따른다고 가정하면 프로빗기반(probit-based) 알고리즘이 도출된다. 로짓기반 확률적 통행배정 알고리즘은 비관련 대안의 독립성(independence of irrelevant alternatives: IIA) 특성과 등분산문제 등으로 인해 중복 및 상관된 경로(overlapping or correlated path)와 총통행거리 별 단위효용변화에 대한 적절한 고려가 불가능하다는 단점이 있다. 프로빗기반 확률적 통행배정 알고리즘은 이러한 로짓기반 확률적 통행배정 알고리즘의 약점을 완화시키지만 계산과정이 복잡하여 대규모 네트워크에 적용하기 힘든 단점이 있다.

고정효용함수 V 는 전통적으로 경로의 서비스수준을 설명하는 변수, 통행자의 사회·경제적인 측면을 설명하는 변수 등을 사용하여 추정해왔다.

$$V = \sum \beta_i x_i$$

β_i : i 번째 속성변수의 파라미터

x_i : i 번째 속성변수

경로의 서비스수준을 설명하는 속성변수에는 차내시간, 통행거리, 운임, 운행빈도, 환승횟수, 정차역 수 등이 있으며, 통행자의 사회·경제적인 측면을 설명하는 속성변수에는 소득, 나이, 성별 등이 있다. 이 밖에도 통행자의 경로선택에 영향을 미치는 경로 및 통행자의 속성변수는 상당히 많다. 심지어 Raveau et al.(2011)의 연구에서는 지하철 노선의 토폴로지(topology, 예: 노선의 굴곡도 등)도 경로선택에 영향을 미치는 유의한 변수로 보고되었다. <표 II-6>은 선행연구에서 대중교통 이용자의 경로선택에 영향을 미치는 속성변수와 파라미터를 추정한 결과를 정리한 것이다.

<표 II-6> 효용함수(또는 일반화비용) 변수 및 파라미터 추정 결과

논문명	구분	도보 시간 (분)	대기 시간 (분)	차내 시간 (분)	요금 (백원)	환승횟 수 (회)	환승시 간 (분)	에스컬 레이터 유무 (무:0, 유:1)	수평이 동거리 (100m)	상향 계단수 (100개)	하향 계단수 (100개)	에스컬 레이터 이동시 간(분)	상수
손상훈 외(2007) ¹⁾	서울시내간	-0.27 75	-0.298 0	-0.162 7	-0.004 6	-	-0.222 9	-	-	-	-	-	-
	서울경기간	-0.27 03	-0.294 0	-0.154 0	-0.003 3	-	-0.194 7	-	-	-	-	-	-
천승훈(2010)	지하철 간 환승	-	-	-0.056 62	-0.019 8	-0.683 7	-	-	-	-	-	-	1.2332
	지하철 환승없음	-	-	-0.056 62	-0.019 8	-0.683 7	-	-	-	-	-	-	1.7445
양창화·손의영 (2000) ²⁾	SP자료	-	-	-0.184	-	-1.881	-0.313	0.368	-	-	-	-	-
	RP자료	-	-	-0.273	-	-4.227	-0.281	1.186	-	-	-	-	-
이경재(2004)	-	-	-	-0.316 2	-	-3.115 0	-	-	-0.003 8	-0.027 4	-0.018 9	-0.012 3	-
Raveau et al.(2011) ³⁾	-	-0.22 9	-0.121	-0.117	-	-0.420	-	0.457	-	-	-	-	-

1) 환승대기시간은 대기시간에 포함하고 환승에 소요되는 도보시간은 도보시간에 포함함

2) 환승시간은 환승을 위해 다른 노선으로 이동하는 시간과 차량대기시간을 합한 시간을 의미함

3) 표에서 제시한 변수뿐만 아니라 수직통행유무(ascending level change): -0.432; 평균제차율: -1.250; 승차불가
데미: -0.413; 착석가능 데미: 0.106; 네트워크 인지도: 0.030; 노선의 굴곡비용(angular cost): -0.038; 출발지에
가까워짐(turn back to origin): -0.503; 도착지에서 멀어짐(turn away from destination): -0.512 변수들이 유
의하게 나타남

선행연구의 결과를 통해서 경로선택에 영향을 미치는 속성변수의 종
류와 그 파라미터의 추정 결과는 연구자, 지역, 연구시기, 분석한 데이터
형태, 통행특성 별로 달라질 수 있음을 알 수 있다. 이것은 경로선택에
대한 통행자의 효용함수를 추정하는 것이 시간과 비용이 많이 소요되는
어려운 작업임을 시사한다. 또한, 출발지-도착지 별로 이러한 속성변수
의 종류나 영향력이 달라질 수 있을 뿐 아니라, 개별 의사결정자를 대상
으로 추정한 효용함수는 시·공간적 범위에 따라 달라질 수 있으므로 범
용적 적용을 위해서는 충분히 많은 조사와 연구가 수반되어야 한다는 한
계가 있다.

경로기반 확률적 알고리즘이 갖는 두 번째 특징은 출발지에서 도착지까지 각 경로의 선택확률을 구할 때 링크와 노드를 분석단위로 살펴보는 것이 아니라 링크의 집합으로 구성되는 경로 전체를 대상으로 한다는 점이다. 즉, 링크기반 통행배정 알고리즘이 링크와 노드 정보만을 필요로 하는 것과는 달리, 경로기반 통행배정 알고리즘은 출발지에서 도착지까지 존재하는 모든 경로 각각에 대한 정보가 필요하다. 그러나 많은 수의 링크와 노드로 구성된 대규모 네트워크에서는 출발지와 도착지 간에 무수히 많은 경로가 존재하기 때문에 모든 경로를 나열(enumeration)하는 것은 사실상 불가능하다. 소규모의 네트워크에서도 모든 경로를 나열하는 것은 쉽지 않으며 계산적으로 비효율적이다. 또한 통행자들이 경로를 선택할 때 고려하지 않는 경로에도 선택확률을 배분함으로써 오차가 발생한다.

한편, 경로나열문제를 해결하기 위한 노력으로 전체 경로집합에서 일부 경로만을 추출하여 선택집합으로 사용하는 방법론들이 제시되어 왔다. 그 중에서 가장 많이 적용되고 있는 방법은 k 개의 최단경로를 찾아서 경로 선택집합(choice sets)으로 구축하는 k -최단경로 알고리즘(Yen, 1971; Azevedo et al. 1993; Eppstein, 1999 등)이 있다. 하지만 경로선택집합을 일괄적으로 k 개로 설정하는 것은 다분히 임의적이다. 이와 관련하여 Bekhor et al.(2008)는 경로 선택집합의 크기와 경로선택 방법론에 따라 통행배정 결과가 달라지는 것을 보였는데, 이 결과를 통해서 경로선택집합에 속한 대안경로 수를 일괄적으로 k 개로 설정하는 것은 네트워크의 규모와 형태(topology)에 따라 선택대안 경로의 수를 달리 하지 못한다는 점에서 비합리적이다. 한편, 신성일 외(2005, 2010)와 천승훈(2010)에서는 최적의 경로와 비교하여 비용측면에서 크게 차이가 나지 않는 경로 또는 비합리적이라 판단되는 경로를 제외한 경로를 ‘유사경로’라 정의하고 유사경로만을 선택대안집합으로 설정하였다. 그런데 유사경로 결정비율에 대한 타당한 근거는 제시되지 않고, 적합한 수준의 결정비율을 구하기 위해서는 통행행태조사가 필요해 보인다.

4. 선행연구의 시사점

대중교통 통행배정 알고리즘은 전통적인 승용차 통행배정 알고리즘의 분류방법과 마찬가지로 통행자의 인지오차 고려여부에 따라 크게 결정적 통행배정 알고리즘과 확률적 통행배정 알고리즘으로 분류할 수 있다. 또한, 분석단위가 링크단위인지 경로단위인지에 따라 다시 링크기반 통행배정 알고리즘과 경로기반 통행배정 알고리즘으로 나뉘며 최종적으로 <표 II-2>와 같이 분류할 수 있다.

대중교통 통행배정에 주로 사용되고 있는 알고리즘은 링크기반 결정적 통행배정 알고리즘과 경로기반 확률적 통행배정 알고리즘이다. 그런데 두 알고리즘은 각각 통행자의 경로선택에 대한 ‘결정론적’인 가정, 분석단위가 ‘경로기반’이라는 측면에서 결함이 나타난다.

먼저 링크기반 결정적 알고리즘에서 통행자의 경로선택에 대한 ‘결정론적’인 가정이 갖는 결함을 Spiess and Florian(1989)의 최적전략 알고리즘을 통해 살펴보았다. 최적전략 알고리즘은 대중교통 이용자가 여러 서비스 노선(common lines)이 통과하는 정차역에서 대기하면서 자신의 총 통행시간을 최소화하기 위한 서비스 노선을 선택할 수 있다는 대중교통 이용자의 통행특성을 설명하기 위하여 최적전략이라는 개념을 도입하였다. 최적전략이란 총‘예상’통행시간을 최소화하는 전략의 집합으로 정의되며 모든 통행은 최적전략에 전량 배정된다. 최적전략 알고리즘은 이용확률을 갖는 다중경로 집합에 통행배정할 수 있지만, 여기서 경로의 이용확률은 효용극대화이론을 기반으로하는 이용확률과는 다른, 분석가의 관점에서 정의하는 결정론적인 측면의 이용 ‘비율’이다. 즉, 최적전략 알고리즘은 통행배정 결과가 다중경로로 나타날 수 있지만, 모든 통행자가 네트워크의 교통상황을 정확히 알고 있으며, 경로선택의 기준이 차이가 없다고 가정함으로써 최적전략에 전량 통행배정하는 한계점이 있다. 또한 각 링크(전략)의 이용확률은 최적전략에 포함된 노선의 운행빈도 비율의 함수로서 다른 서비스 특성보다 운행빈도에 더 큰 가중치를 두며

다른 서비스수준과는 독립적이다. 따라서 대기시간 및 환승시간 가중치가 증가하여 총통행비용이 증가하더라도 최적전략이 변화가 없다면 링크의 이용확률은 운행빈도 비율로 고정되어 나타나는 비합리적인 결과가 도출된다. 이러한 한계점을 극복하지 못한 채, 승차시간, 승차시간 가중치, 대기시간 가중치, 차두간격에 대한 대기시간 비율, 접근시간 가중치 등 다수의 변수를 사용하여 대중교통 이용자의 통행행태를 설명하려고 하지만, 오히려 다수의 파라미터를 추정하는데 많은 시간 및 비용이 소요되고, 분석의 시·공간적 범위가 다를 경우 범용적으로 적용할 수 없어서 충분한 사전연구가 수반되어야만 한다는 어려움이 있다.

두 번째로, 경로기반 확률적 알고리즘은 통행자들이 같은 경로를 사용하더라도 그 통행시간을 다르게 인지한다는 ‘확률론적’인 가정을 함으로써 앞서 설명한 링크기반 결정적 알고리즘의 ‘결정론적’인 가정을 완화하였다. 그러나 경로기반 확률적 알고리즘은 분석단위가 ‘경로’이기 때문에 분석하는 출발지-도착지 사이에 존재하는 모든 경로의 정보가 필요하다. 많은 수의 링크와 노드로 구성된 대규모 네트워크에서는 출발지와 도착지 간에 무수히 많은 경로가 존재하기 때문에 모든 경로를 나열(enumeration)하는 것은 사실상 불가능하다. 또한 통행자들이 경로를 선택할 때 고려하지 않는 경로에도 선택확률을 배분함으로써 그만큼의 오차가 발생할 수 있다. 경로나열문제를 해결하기 위한 노력으로 전체 경로집합에서 일부 경로만을 추출하는 방법론들이 제시되어 왔지만 경로선택집합을 설정하는 기준이 임의적이다.

또한, 경로기반 확률적 알고리즘은 통행자가 각 경로에 대한 효용을 갖고, 효용극대화의 원칙에 따라 가장 높은 효용을 갖는 경로를 선택한다고 가정한다. 따라서 경로의 효용함수를 구축하여야 하는데, 경로의 효용함수에 포함될 수 있는 속성변수는 상당히 많다. 전통적으로 경로의 서비스수준을 설명하는 변수, 통행자의 사회·경제적인 측면을 설명하는 변수 등을 사용하여 추정해왔는데, 이 밖에도 통행자의 경로선택에 영향을 미치는 변수들이 계속적으로 논의되고 있다. 이처럼 경로기반 통행배

정 알고리즘은 출발지-도착지 간 경로를 나열하는 문제와 더불어 각 경로의 효용을 설명하는 속성변수를 찾고, 각 변수의 영향력을 추정해야 한다는 난제를 해결해야 한다. 효용함수를 추정하더라도 출발지-도착지 별로 경로의 속성변수의 종류나 영향력이 달라질 수 있으므로 범용적 적용에 대한 충분한 조사와 연구가 수반되어야만 한다.

지금까지의 선행연구 고찰을 통해서 통행자의 경로선택에 대한 가정이 ‘결정론적’인 알고리즘과 분석단위가 ‘경로’인 알고리즘의 한계점을 살펴보았다. 본 연구에서는 두 알고리즘의 한계점을 극복하기 위한 대안으로 ‘링크’기반의 ‘확률적’인 가정을 전제로 하고, 아울러 많은 변수를 요구하지 않으면서도 실제 통행패턴을 잘 묘사하며, 이례적인 통행패턴이 발생하는 구간에서도 범용적으로 유연하게 적용할 수 있는 알고리즘에 대하여 논의해 보고자 한다.

III. 링크기반 확률적 대중교통 통행배정 알고리즘

1. 개요

분석단위가 ‘경로’인 경로기반 통행배정 알고리즘과 통행자의 경로선택에 대한 ‘결정론적’ 가정을 전제로 하는 결정적 통행배정 알고리즘의 한계점을 극복하기 위한 대안으로 링크기반 확률적 통행배정 알고리즘을 제안할 수 있다. 링크기반 확률적 통행배정 알고리즘은 분석단위가 ‘링크’이고 통행자에 대한 가정이 ‘확률적’이라는 두 가지 특성을 갖는다. 먼저, 링크기반 통행배정 알고리즘은 경로기반 통행배정 알고리즘이 링크의 집합으로 구성되는 경로 전체를 분석단위로 하여 모든 경로 각각에 대한 정보를 요구했던 것과는 달리, 링크와 노드 변수만을 분석단위로 하기 때문에 네트워크를 구현하는 정도의 정보만을 요구한다. 둘째, 링크기반 확률적 통행배정 알고리즘은 통행자가 같은 경로를 사용하더라도 그 통행시간을 다르게 인지하는 것을 반영하기 위해 통행시간을 단일 값이 아닌 확률분포로 적용한다.

링크기반 확률적 통행배정 알고리즘의 가장 대표적인 예는 Dial(1971)의 알고리즘이다. 승용차의 통행배정을 목적으로 고안된 Dial 알고리즘은 알고리즘의 우수성에도 불구하고 대중교통 통행배정에의 적용 가능성에 대한 논의는 일부만이 이루어졌다. 예를 들어, Nguyen et al.(1998)은 전략(strategy)과 하이퍼패스(hyperpath)의 틀(framework)에서 Dial 알고리즘의 효율적 경로 정의를 응용한 ‘efficient hyperarc’의 개념을 도입하고, 링크(hyperarc)를 기반으로 분석함으로써 경로나열문제를 해결하고자 하였다. 그러나 이 접근법은 도시철도 이용자의 통행행태 분석에는 이론적으로 적합하지 않은 한계점이 드러난다. 전략의 개념과 하이퍼패스란 틀은 버스정류장(stop)을 이용하는 이용자의 통행행태를 설명하기 위해 고안된 것으로 경쟁노선 집합에 대한 버스 승객의 선택은 합리적으로 묘사할 수 있지만, 각 노선이 공간적으로 분리된 지하철 정거장(station)을

이용하는 통행자의 경로선택을 설명하는 데에는 적합하지 않다. 한편, 국내에서 대중교통 통행배정에 Dial 알고리즘을 적용한 연구는 이재섭·김익기(2001)가 있다. 이재섭·김익기(2001)는 Dial 알고리즘과 상용 프로그램인 TRANPLAN, EMME/2의 결정적 대중교통 통행배정 알고리즘을 비교하고, Dial 알고리즘이 이론적 배경 및 분석의 융통성 측면에서 더 우수하다고 주장하였다. 그러나 Sioux Falls 네트워크에 정산되지 않은 임의의 θ 값을 적용하였으며, 이를 통해 시사점만을 도출하였다. 또한 동 연구에서 전제하는 대중교통은 암묵적으로 버스를 지칭하는 것으로 판단된다.

다음의 2절에서는 Dial 알고리즘에 대한 기본적인 내용을 고찰하고, 3절에서 도시철도 통행배정 분석에 Dial 알고리즘을 적용하기 위한 네트워크 표현방식을 제안한다. 또한 전통적 Dial 알고리즘으로는 분석하기 어려운 출발지-도착지 별 링크(또는 경로)의 이용확률 산정 방안을 살펴본다. 마지막 4절에서는 통행자의 인지오차 수준을 조절할 수 있는 파라미터 θ 와 그 정산 방안에 대하여 살펴본다.

2. Dial 알고리즘

1) Dial 알고리즘의 개요

Dial 알고리즘은 승용차의 통행배정을 목적으로 고안된 로짓기반의 확률적 네트워크 부하 알고리즘으로서 파라미터 θ 를 통해 이용자의 인지도차 수준을 조절할 수 있다. Dial 알고리즘은 두 단계를 거쳐 네트워크에 통행량을 부하하는데, 먼저 전방 진행(forward pass) 단계에서 링크의 이용 가중치(link weights)을 구하고, 두 번째 단계인 후방 진행(backward pass) 단계에서 링크의 이용확률에 따라 효율적 경로(efficient paths)에 통행량을 배정한다. 효율적 경로는 두 가지 방법으로 정의된다. 첫 번째는 출발지로부터 멀어지고, 도착지로 가까워지는 링크들로 구성된 경로로 정의되고, 두 번째는 출발지로부터 멀어지는 링크들로 구성된 경로로 정의된다. 두 번째 효율적 경로의 정의는 첫 번째 정의보다 완화된 기준으로서, 이 정의를 적용하면 고려해야 할 효율적 경로의 수는 늘어나지만 도착지로부터의 거리를 고려하지 않아도 되어 한 번의 알고리즘 실행으로 한 개의 출발지에서 모든 도착지에 대한 통행배정을 완료할 수 있다. 두 가지 정의 중 어느 정의를 적용하더라도 효율적 경로에 포함되지 않는 링크의 이용확률은 0이 되며 통행이 배정되지 않는다. 효율적 경로의 정의는 네트워크에서 이용되지 않을 것으로 판단되는 비합리적인 링크를 효과적으로 제거하고, 특히 루프형태의 경로가 제외되어 비상식적인 순환통행(cyclic flow)을 방지한다.

본 연구에서 살펴보는 Dial 알고리즘은 두 번째 효율적 경로의 정의를 적용한다.

0. 초기화 단계: 각 노드 i 에 대하여 다음의 정보를 입력한다.

- i. $y(i)$: 출발지(O)에서 출발하여 노드 i 에서 종결하는 통행량
- ii. $p(i)$: 출발지(O)에서 노드 i 까지의 최단경로거리
- iii. I_i : 노드 i 에서 시작하는 모든 링크의 집합
- iv. F_i : 노드 i 에서 종결하는 모든 링크의 집합

이 후, 출발지(O)의 노드가중치 $W(O)=1$ 를 제외한 모든 노드의 가중치를 0으로 초기화하고, 링크의 우도(likelihood, $l(i, j)$)를 계산한다.

$$W(i) = \begin{cases} 1 & \text{if } i = O \\ 0 & \text{otherwise} \end{cases}$$

$$l(i, j) = \begin{cases} \exp\theta[p(j)-p(i)-t(i, j)] & \text{if } p(i) < p(j) \\ 0 & \text{otherwise} \end{cases}$$

여기서, θ 는 인지오차 파라미터이고 $t(i, j)$ 는 링크(i, j)의 길이를 의미한다.

1. 전방진행 단계: 한 번도 탐색하지 않은 노드 중에서 출발지로부터 가까운 노드 i , 즉, $p(i)$ 가 작은 노드 i 부터 아래의 링크 가중치(w)와 노드 가중치(W)를 계산한다. 모든 노드를 탐색했으면 후방진행 단계로 넘어간다.

$$w(i, j) = W(i) \times l(i, j)$$

$$W(j) = \sum_i w(i, j)$$

2. 후방진행 단계: 한 번도 탐색하지 않은 노드 중에서 출발지로부터
 먼 노드 i , 즉, $p(i)$ 가 큰 노드 i 부터 링크의 통행
 량 $x(i, j)$ 을 구한다. 출발지(O)까지 탐색하면 통행
 배정을 완료한다.

$$x(i, j) = y(j) \times \frac{w(i, j)}{W(j)}$$

$$y(i) \leftarrow y(i) + x(i, j)$$

Dial 알고리즘은 확률적 통행배정 방식이라는 점 이외에도 분석단위가
 링크기반이라는 특성을 갖는다. 위의 알고리즘의 설명에서 확인할 수 있듯
 이, Dial 알고리즘은 링크와 노드 변수만을 사용하여 통행배정량을 계산한
 다. 따라서 경로기반 알고리즘이 갖는 경로나열(path enumeration) 문제를
 해결하는 계산적으로 효율적인 알고리즘이다. 또한, Dial 알고리즘의 통행
 배정 결과는 효율적 경로만으로 구축된 축소 네트워크(reduced network)
 를 동일하게 적용한 경로 및 로짓기반 알고리즘의 통행배정 결과와 동등
 하게 나타난다. 이는 Dial(1971), Sheffi(1985)가 다음과 같이 증명하였다.

경로 P 를 이용할 확률은 경로 P 를 구성하는 링크의 우도 $l(i, j)$ 의
 곱에 비례한다. 즉,

$$\Pr(P) = k \prod_{(i, j) \in P} l(i, j).$$

위 식에 $l(i, j) = \exp \theta [p(j) - p(i) - t(i, j)]$ 를 대입하여 풀면,

$$\begin{aligned} \Pr(P) &= k \prod_{(i, j) \in P} \exp \theta [p(j) - p(i) - t(i, j)] \\ &= k \exp \theta \sum_{(i, j) \in P} [p(j) - p(i) - t(i, j)] \\ &= k \exp \theta \left[p(d) - \sum_{(i, j) \in P} t(i, j) \right]. \end{aligned}$$

$\sum \Pr(P) = 1$ 이므로,

$$k = \frac{1}{\sum \exp \theta \left[p(d) - \sum_{(i,j) \in P} t(i,j) \right]}.$$

따라서,

$$\Pr(P) = \frac{\exp \theta \left[p(d) - \sum_{(i,j) \in P} t(i,j) \right]}{\sum \exp \theta \left[p(d) - \sum_{(i,j) \in P} t(i,j) \right]}.$$

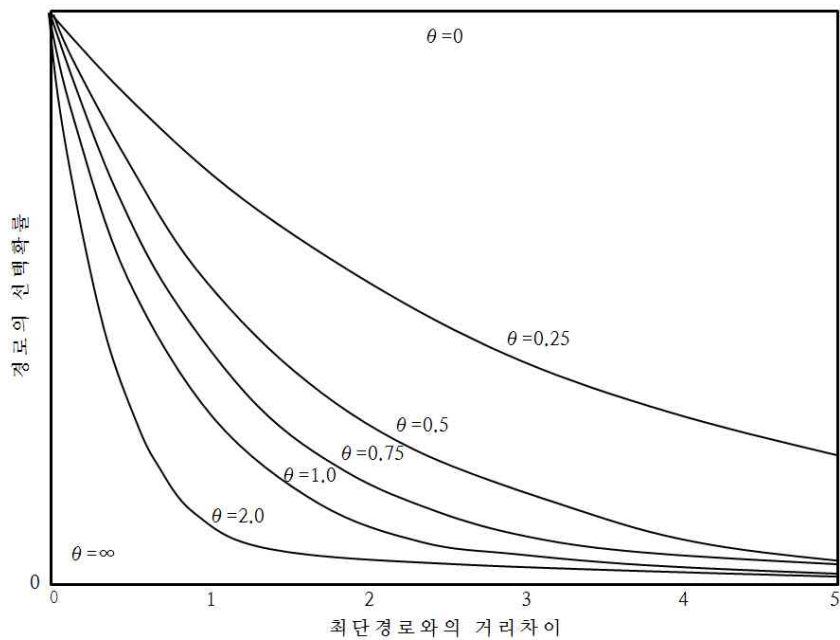
위의 식은 경로 및 로짓기반의 확률적 통행배정 모형에서 경로선택 확률을 구하는 식과 형태가 동일하다. 이처럼, Dial 알고리즘은 경로기반 통행배정 알고리즘의 경로나열 문제를 해결하면서도 그 결과는 동일한 네트워크를 사용할 경우 경로기반 통행배정 결과와 동일하다.

지금까지의 내용을 정리하면, Dial 알고리즘은 최적전략 알고리즘과 같은 결정적 통행배정 알고리즘이 통행자의 인지오차를 고려하지 못하는 결함을 해결한다. 또한, Dial 알고리즘은 링크와 노드 정보만을 요구하여 경로나열문제를 해결할 뿐만 아니라, 비합리적인 경로를 효율적 경로의 정의를 통해 사전에 제거하기 때문에 계산적으로 효율적이다. 게다가 적용하는 네트워크가 동일하다면 링크기반의 Dial 알고리즘의 통행배정 결과는 경로를 분석단위로 하여 구한 통행배정 결과와 동일하다.

2) 인지오차와 파라미터 θ

모든 통행자가 항상 최단경로를 정확하게 인지하는 것은 아니다. 왜냐하면, 대안경로 간의 통행시간 차이가 충분히 작아서 통행자가 이를 인지하지 못할 수도 있고, 혹은 과거의 통행에서 경험한 이례적인 통행시간이 경로선택에 영향을 미칠 수도 있다. 또는 통행자 별로 경로를 선

택할 때 중요하게 여기는 경로의 속성이 다르기 때문에 최단경로가 아닌 다른 경로가 이용될 수도 있다. II장 3절에서 살펴보았던 경로기반 확률적 알고리즘은 통행자의 인지오차에서 비롯되는 경로선택 패턴을 다양한 변수로 구축한 효용함수를 통해 설명했지만, Dial 알고리즘은 통행거리(또는 시간)를 유일한 변수로 사용하고 파라미터 θ 를 통해 인지오차를 조절한다. 예컨대, θ 값이 커지면 통행자 간 편차(경로의 속성을 인지하는 편차)와 통행자의 통행시간에 대한 인지오차가 줄어들어 최단경로를 이용할 확률이 커지고, 반대로 θ 값이 작아지면 통행자 간 편차와 통행시간에 대한 인지오차가 커져서 최단경로가 아닌 경로를 선택할 확률이 높아진다.

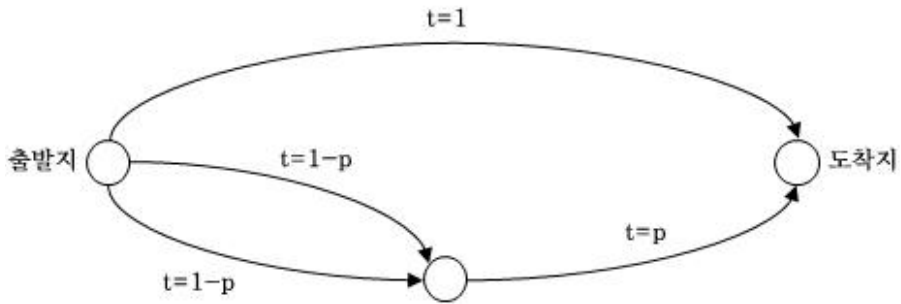


Dial(1971)

<그림 III-1> θ 및 최단경로와의 거리차이 별 경로의 선택확률

3) 로짓기반 알고리즘

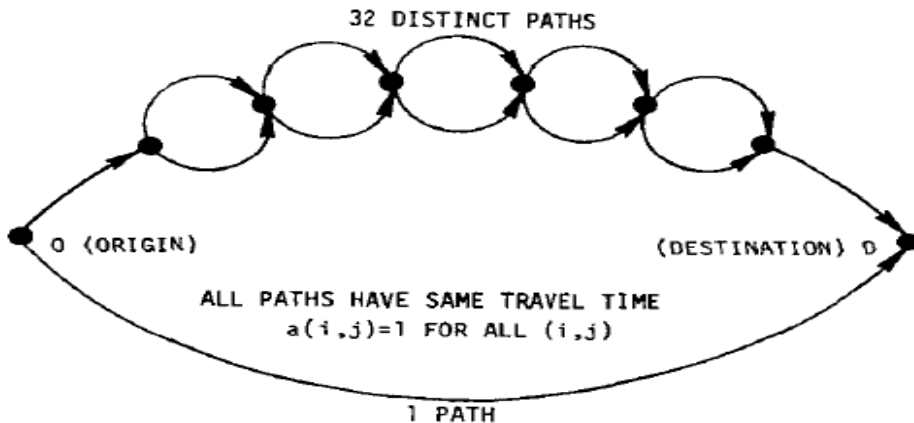
Dial 알고리즘은 확률효용에 대한 분포를 검블(Gumbel)분포로 가정하는 로짓기반(logit-based) 알고리즘이다. 로짓기반 알고리즘의 계산적인 효율성 때문에 Dial 알고리즘은 대규모 네트워크에 적용이 용이하다. 그러나 Dial 알고리즘은 로짓모형의 오차항이 독립적이고 동일하다는(identically and independently distributed: IID) 가정 때문에 두 가지 측면에서 비현실적인 통행배정 결과가 나타날 수 있다. 첫 번째 문제점은 오차항에 대한 독립성(independence) 가정으로 인해 대안의 효용 간 상관성(correlation)을 고려하지 못한다는 점이다. 이로 인해 중복 및 상관된 경로(overlapping or correlated path)에 대한 적절한 고려가 불가능하다. Sheffi(1985)는 <그림 III-2>의 예제 네트워크(4)를 통해서 이 문제점을 설명하였다. 예제 네트워크(4)에는 경로 길이가 1로 모두 같은 총 3개의 경로가 있는데, 아래의 두 경로는 p 만큼의 구간이 중복된다. 중복된 구간 p 의 길이가 0에서부터 1에 가까워질수록 실제 통행자들은 아래의 중복된 두 경로를 하나의 경로로 인지하게 되며, 중복되지 않은 위의 경로를 선택할 확률이 $1/3$ 에서 $1/2$ 로 증가하게 된다. 그러나 로짓기반의 통행배정 알고리즘은 중복구간의 길이에 따른 상관성을 반영하지 못하여 중복구간의 길이와 관계없이 모든 경로의 선택확률이 동일하게 $1/3$ 로 나타난다. 따라서 중복경로(overlapping path)에 통행량이 과다 배정되는 비현실적인 결과가 도출될 수 있다.



Sheffi(1985)

<그림 III-2> 예제 네트워크(4)

로짓기반 통행배정 알고리즘이 대안 간 상관성을 고려하지 못해서 중복경로에 과다 통행배정하는 점은 <그림 III-3>의 예제 네트워크(5)에서 확연하게 드러난다. 예제 네트워크(5)에는 출발지-도착지 간 통행길이 같은 총 33개의 경로가 있는데, 로짓기반 통행배정 알고리즘으로 중복되지 않은 경로의 선택확률을 구하면 $1/33(=0.03)$ 로 비현실적인 결과가 도출된다.



Tobin(1977)

<그림 III-3> 예제 네트워크(5)

로짓기반 통행배정 알고리즘의 두 번째 문제점은 오차항이 동일하게 (identically) 분포한다는 가정으로 인해 인지통행시간의 분산을 고정되고 동일한 값으로 반영한다는 점이다. 이를테면, <그림 III-4>의 (a)와 (b)에서 최단경로 통행시간과의 시간차이 5분((a)의 경우 10분-5분, (b)의 경우 125분-120분)은 경로의 총통행시간과 무관하게 경로의 선택확률에 동일한 비율로 작용하여, (a)와 (b)의 통행배정 결과는 동일하게 나타난다. 그러나 실제로는 (b)의 경우와 같이 총통행시간이 길수록 실제 통행시간의 분산이 커질 뿐만 아니라 통행시간에 대한 이용자의 인지오차 역시 커지기 때문에(Herman and Lam(1974)의 결과를 Tobin(1977)에서 재인용) 이 결과는 비합리적이다. Herman and Lam(1974)의 연구에 따르면, 총통행시간이 긴 경로일수록 인지오차의 분산이 크기 때문에 (a)와 (b)처럼 대안경로의 통행시간 차이가 같더라도 (b)의 경우는 통행시간이 더 긴 경로를 선택하는 통행자가 (a)에 비해 상대적으로 더 많아야 한다.



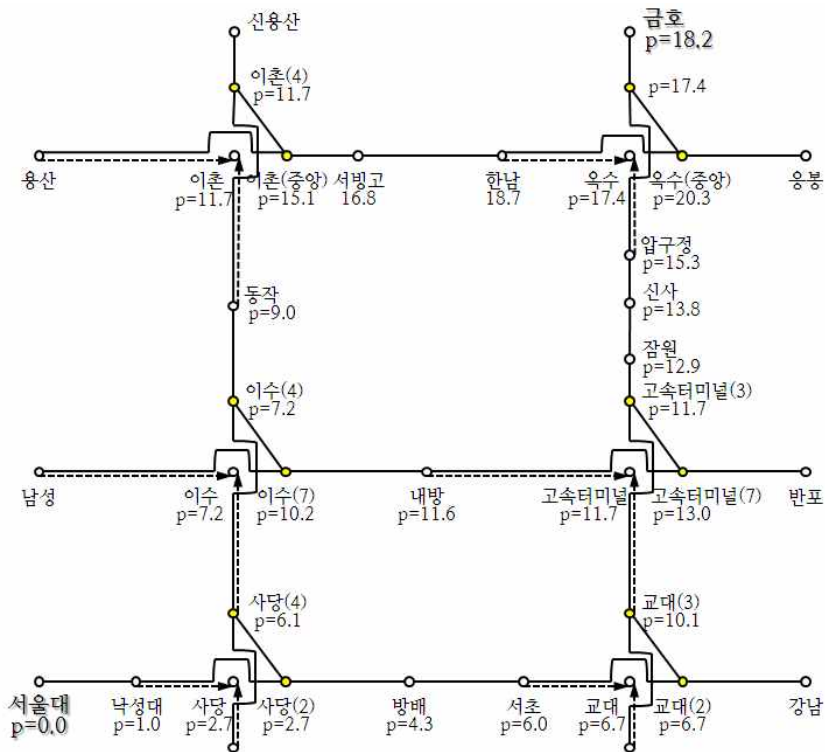
Sheffi(1985)

<그림 III-4> 예제 네트워크(6)

로짓기반 통행배정 알고리즘은 위에서 기술한 문제점에도 불구하고 대규모 네트워크에서도 전량배정기법과 비슷한 수준의 계산적 효율성 때문에 선호되어 왔다. 게다가 위에서 기술한 로짓기반 통행배정 알고리즘의 두 가지 문제점은 일반적으로 나타나는 것은 아니며, 대안 경로의 통행시간이 서로 상관성이 적거나 또는 확률 분포가 유사한 네트워크에서는 문제가 되지 않는다(Sheffi, 1985).

4) 효율적 경로

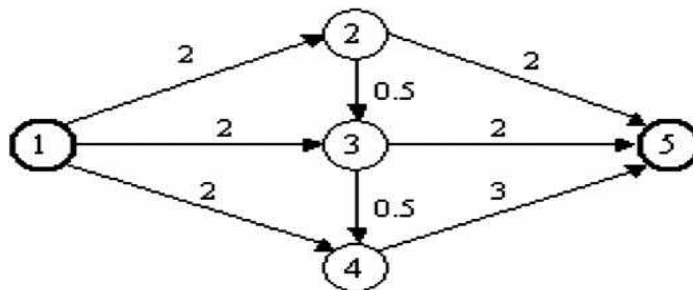
Dial 알고리즘에서 효율적 경로의 정의는 네트워크에서 이용되지 않을 것으로 판단되는 비합리적인 링크를 효과적으로 제거하고, 특히 루프 형태의 경로를 제외시켜 비상식적인 순환통행(cyclic flow)을 방지한다. 예를 들어, <그림 III-5>는 서울대입구-금호 간 전체 네트워크이고 p 는 출발지인 서울대입구로부터의 최단거리를 의미하는데, 출발지로부터 멀어지는, 즉, $p(i) < p(j)$ 인 링크(i, j)로만 구성된 효율적 경로는 <그림 III-6>과 같이 나타난다.



<그림 III-5> 서울대입구-금호 간 전체 네트워크

이처럼 Dial 알고리즘은 효율적 경로의 정의를 통해 지나치게 우회하는 경로를 사전에 효과적으로 제거하고, 실질적으로 이용될 가능성이 높은 경로만으로 축소된 네트워크(reduced network)를 구축함으로써 계산상으로 효율적이고 행태적으로 합리적인 통행배정 결과를 얻을 수 있다.

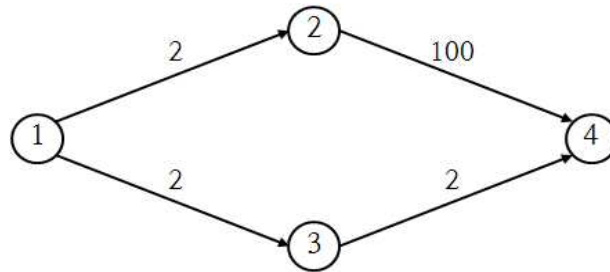
그러나 효율적 경로의 정의는 통행배정되는 경로를 다소 엄격한 기준으로 제한하는데, 이러한 점 때문에 모든 출발지-도착지 쌍에 대하여 합리적인 결과를 도출하지 못할 수도 있다. Akamatsu(1996)는 효율적 경로의 정의로 인해서 총통행비용이 상대적으로 더 작은 경로가 선택대안집합에서 제외되는 비합리적인 결과가 도출될 수 있음을 보였다. <그림 III-7>의 예제 네트워크(7)을 통해 이 문제점을 쉽게 확인할 수 있다. <그림 III-7>에서 노드 ①을 출발지로, 노드 ⑤를 도착지로 하였을 때 출발지-도착지 간 ①→④→⑤ 노드를 통과하는 경로를 A경로라 하고 ①→②→③→⑤ 노드를 통과하는 경로를 B경로라 하자. 경로 A의 총통행거리는 $5(=2+3)$ 이고 경로 B의 총통행거리는 $4.5(=2+0.5+2)$ 인데, Dial 알고리즘의 효율적 경로 정의에 의하면, A경로는 효율적 경로에 포함되어 통행이 배정되는 반면에 B경로는 ②→③ 구간에서 효율적 경로의 정의를 만족하지 않아 통행배정 되지 않는 비합리적인 결과가 도출된다. 이러한 문제는 Dial의 효율적 경로가 전체 경로 통행비용은 고려하지 않고 출발지나 도착지로부터 각 노드의 상대적 위치만을 고려하기 때문이다.



Si et al.(2010)

<그림 III-7> 예제 네트워크(7)

이와 같이, 효율적 경로의 정의가 출발지나 도착지로부터 각 노드의 상대적 위치만을 고려하기 때문에 발생할 수 있는 두 가지 추가적인 문제점이 있다. 첫째, 링크의 길이를 고려하지 못한다는 것이고 둘째, 알고리즘은 확실적인 가정을 전제하지만 효율적 경로의 정의는 결정론적이라는 점이다. 이를 설명하기 위해 <그림 III-8>과 같은 다소 과장된 네트워크를 예로 들 수 있다.



<그림 III-8> 예제 네트워크(8)

<그림 III-8>의 네트워크에서 출발지와 도착지가 각각 노드 ①, 노드 ④이고 노드 ②와 노드 ④를 연결하는 링크의 길이는 100으로 상당히 큰 값이다. 경로 ①→③→④의 총통행거리가 4라는 점을 감안한다면 경로 ①→②→④는 상당히 우회하는 경로이다. 그러나 효율적 경로의 정의는 링크의 길이는 고려하지 않고 출발지에서 각 노드까지의 최단거리만을 고려하기 때문에 경로 ①→②→④ 역시 효율적 경로에 포함된다.

또한, 효율적 경로의 정의는 네트워크에 대한 정보를 모든 이용자가 정확하게 인지해야만 하는 결정론적인 측면있기 때문에 확실적 알고리즘인 Dial 알고리즘과 모순된다.

이러한 두 가지 측면에서의 문제점을 해결하기 위하여 Tobin(1977)은 경로식별계수(coefficient of discrimination) δ 를 도입하여 효율적 경로의 정의를 수정하였다. 경로식별계수는 통행자가 총통행거리의 수준에 따라 최단경로와 대안경로의 통행시간 차이를 인지하지 못하는 정도를 의미하는 것으로 기본적인 개념은 다음과 같다.

$$t(P)-t(P') \leq \delta g(P)$$

여기서 $t(P)$ 는 경로 P 의 통행시간, $t(P')$ 는 최단경로 P' 의 통행시간, $g(P)$ 는 경로 P 의 속성함수이다. 즉, 주어진 δ 값에 대해 위 식을 만족하는 경로 P 는 통행자에게 최단경로와 다르게 인지되지 않는다는 것이다. 이러한 특성을 반영하여 효율적 경로의 정의를 다시 나타내면 다음과 같다.

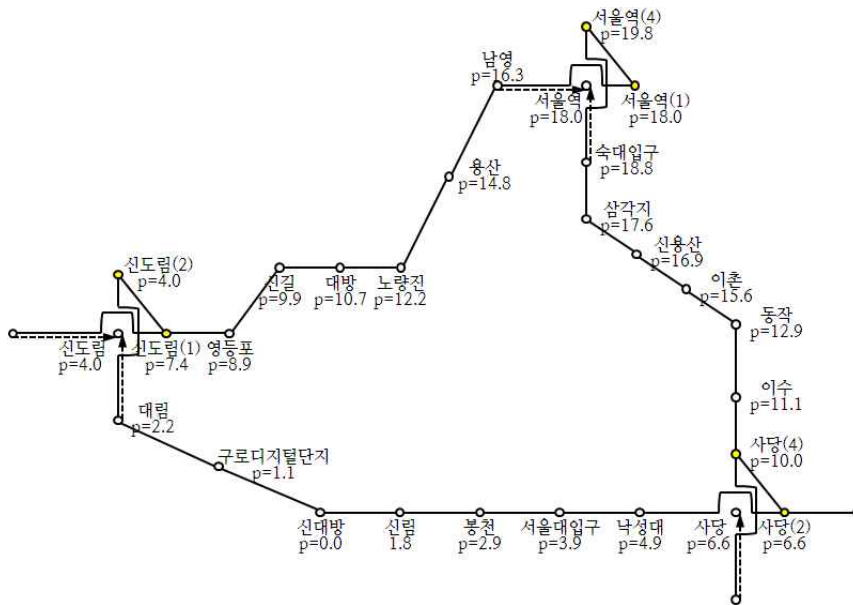
$$p(i)+t(i,j)-p(j) \leq \delta t(i,j) \text{ and } p(i) < p(j)$$

한편, Leurent(1997)는 위의 정의를

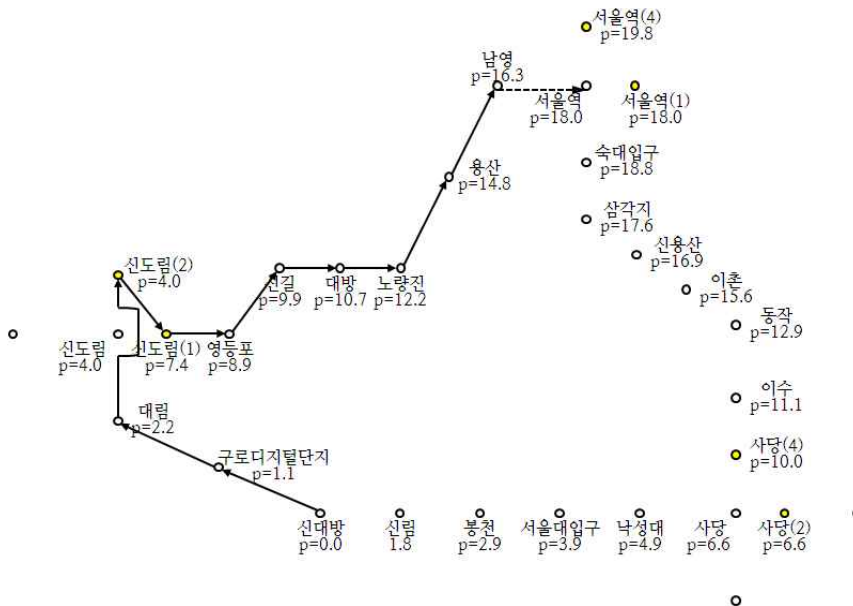
$$t(i,j) \leq (1+h_d^a)\{p(j)-p(i)\} \text{ and } p(i) < p(j)$$

로 다시 쓰고, h_d^a 를 각 도착지 d 에 대한 링크 a 의 팽창율(elongation ratio)이라 하였다. 이 때 h^a 는 도시 간에서는 1.6(USAP(1992)의 결과를 Leurent(1997)에서 재인용)을, 도시 내에서는 1.3~1.5의 값(Tagliacozzo and Pirzio(1973)의 결과를 Leurent(1997)에서 재인용)을 적용하였다.

그러나 Tobin(1977)과 Leurent(1997)에서도 역시 루프형태의 네트워크로 인해 비상식적인 순환통행(cyclic flow)이 발생하는 것을 방지하기 위하여 $p(i) < p(j)$ 라는 다소 엄격한 기준을 효율적 경로의 정의에 포함시켰다. 이로 인해, 실제 이용될 가능성이 높은 경로도 선택대안집합에서 제외될 수 있는데, 실제 수도권 도시철도 네트워크의 일부 출발지-도착지 구간에서도 이러한 문제점이 나타난다. 예를 들어, 신대방-서울역 간의 실제 이용된 경로와 효율적 경로가 다르게 나타나는 점을 살펴볼 수 있다. 신대방-서울역 간 전체 경로와 효율적 경로는 각각 <그림 III-9>, <그림 III-10>과 같다.



<그림 III-9> 신대방-서울역 간 전체 경로



<그림 III-10> 신대방-서울역 간 효율적 경로

<표 III-2>는 신대방-서울역 간 전체 경로와 효율적 경로를 정리한 것이다.

<표 III-2> 신대방-서울역 간 전체 경로 및 효율적 경로

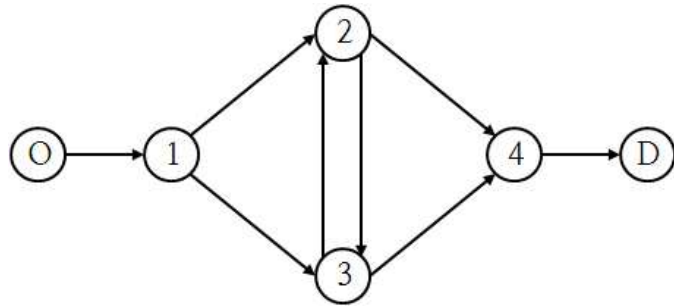
구분	경로	차내거리(km)	환승횟수(회)	효율적 경로
경로 A	신대방-신도림-서울역	14.6	1	○
경로 B	신대방-사당-서울역	16.4	1	×

경로 B 역시 이용될 가능성이 높아보이지만 경로 A만이 효율적 경로에 포함되는 사실을 확인할 수 있다.

Bell(1995)과 Akamatsu(1996)은 효율적 경로의 정의로 인한 문제점들을 효율적 경로의 정의를 수정함으로써 해결하기보다는 배정되는 경로를 제한하지 않는 방법, 즉, 비효율적인 경로에도 통행배정을 하는 방법으로 해결하고자 하였다. 효율적 경로의 정의없이 Dial 알고리즘을 적용하는 것은 위에서 기술한 문제점들을 해결할 뿐만 아니라, Dial의 부하방식을 SUE 알고리즘으로 발전시킬 때 더 용이하다는 장점도 있다. 효율적 경로를 정의한 SUE 알고리즘은 통행시간을 갱신하는 때 반복마다 효율적 경로에 속하는 링크집합이 달라져서 통행배정 결과가 수렴하지 않을 수 있기 때문이다(Maher, 1998). 그러나 효율적 경로의 정의가 없으면 우회적인 경로나 순환적인(cyclic) 경로에도 통행배정될 수 있다. 예를 들어, <그림 III-11>과 같은 네트워크에서 O-D 간 다음과 같은 경로들에도 통행배정된다.

①→②→③→②→④, ①→②→③→②→③→④, ..., ①→②→③→...→②→③→④.

이러한 순환통행(cyclic flows)으로 인해 링크통행량이 과대 추정될 수 있으며 특히 ②→③ 링크의 길이가 다른 링크의 길이에 비해 작을 때 이 문제는 부각된다.



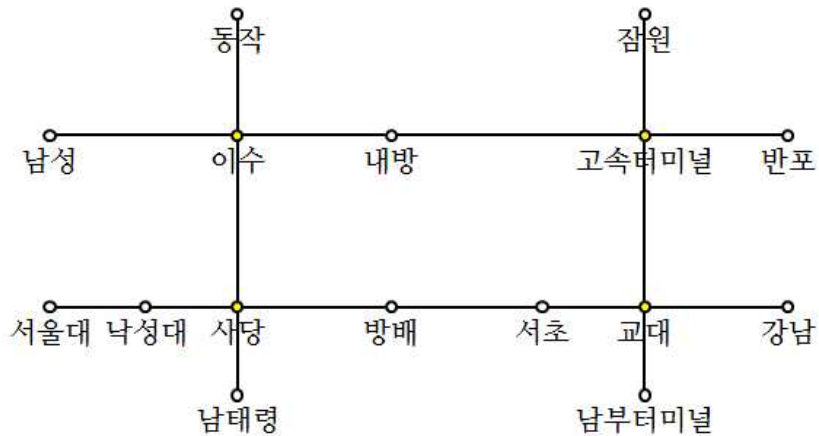
<그림 III-11> 예제 네트워크(9)

한편, Dial 알고리즘의 효율적 경로 정의가 통행량을 부하하는 방식과는 직접적으로 관련된 것이 아니기 때문에(Daganzo and Sheffi(1977)) 본 연구는 효율적 경로에 대한 정의를 수정하지 않고 Dial 알고리즘을 적용한다.

3. Dial 알고리즘의 수정

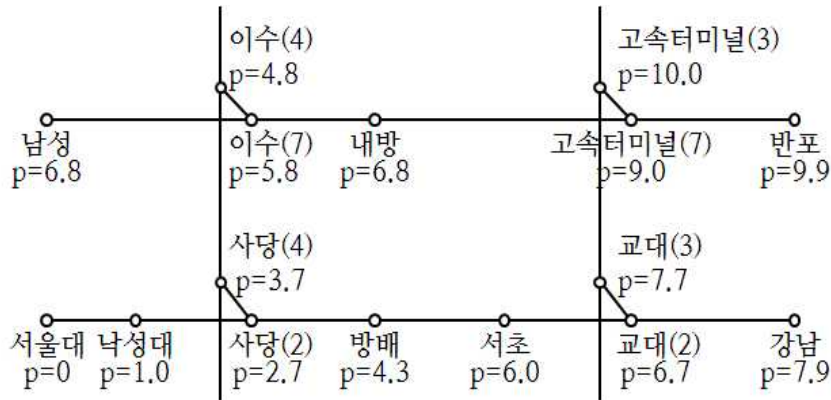
1) 네트워크 표현

통행배정 알고리즘의 적용을 위해서는 우선 네트워크의 표현방식 (representation)을 결정해야 한다. Dial 알고리즘은 승용차 통행배정을 위해 제안된 모형으로서 승용차 통행배정에 적용되는 <그림 III-12>와 같은 네트워크를 그대로 사용할 경우 대중교통 통행에서 특징적으로 나타나는 환승통행을 네트워크 차원에서 반영할 수 없다.



<그림 III-12> 기본적인 네트워크 표현 방식

네트워크에 환승통행을 반영하기 위한 가장 간단한 방법은 <그림 III-13>과 같을 수 있다. 환승역에서 환승 노드를 추가하고, 환승링크를 이어줌으로서 환승링크가 갖는 통행비용 크기만큼 총통행비용에 반영할 수 있다. 예를 들어, 낙성대에서 출발하여 내방까지 통행할 경우 총통행비용은 낙성대-사당(2)-사당(4)-이수(4)-이수(7)-내방의 링크 통행비용의 합으로 산출되며, 두 번의 환승통행비용인 사당(2)-사당(4), 이수(4)-이수(7)이 반영된다.



<그림 III-13> 환승통행 반영 네트워크

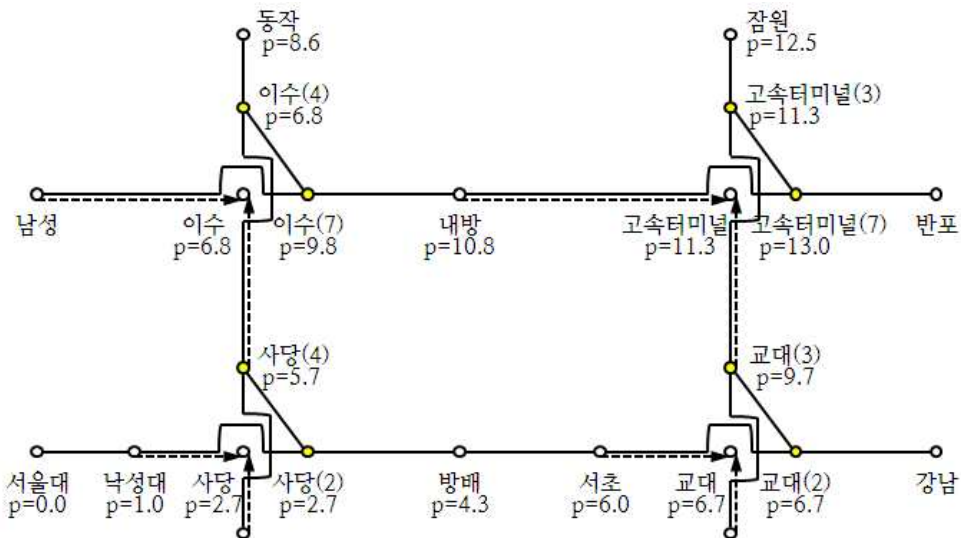
하지만 <그림 III-13>과 같은 네트워크 표현방식에 전통적인 Dial 알고리즘에 적용하면 도착지가 환승역일 경우 모든 선택가능한 경로를 분석할 수 없다. 왜냐하면 위의 표현방식에서 환승역은 하나의 노드로 표현되지 않고, 또한 효율적 경로의 정의에 의해 일부 경로의 선택확률이 0으로 나타나기 때문이다. 예를 들어, 서울대입구역-고속터미널역까지의 두 경로, 경로 A(서울대입구-사당(2)-사당(4)-이수(4)-이수(7)-내방-고속터미널(3 또는 7))와 경로 B(서울대입구-교대(2)-교대(3)-고속터미널(3 또는 7))에 대하여 경로분석을 한다고 가정하자. <그림 III-13>의 네트워크 표현 방식을 따르면 도착지를 고속터미널(3) 또는 고속터미널(7)로 선택하여 분석해야 한다. 만일 고속터미널(7)을 도착지로 선택하여 분석한다면,

$$p(\text{고속터미널}(3)) = 10.0 > p(\text{고속터미널}(7)) = 9.0$$

이다. 여기서 p는 서울대입구역으로부터 각 역까지의 최단경로거리를 의미하며, 환승링크의 길이는 1로, 역간 링크의 길이는 역간거리를 이용하였다. 따라서 효율적 경로의 정의에 의해 고속터미널(3)→고속터미널(7) 링크의 이용확률이 0이 된다. 즉, 경로 B의 선택확률은 0이 되고 경로 A에 전량 배정되게 됨으로써 서울대입구역-고속터미널 간 이용된 경로는

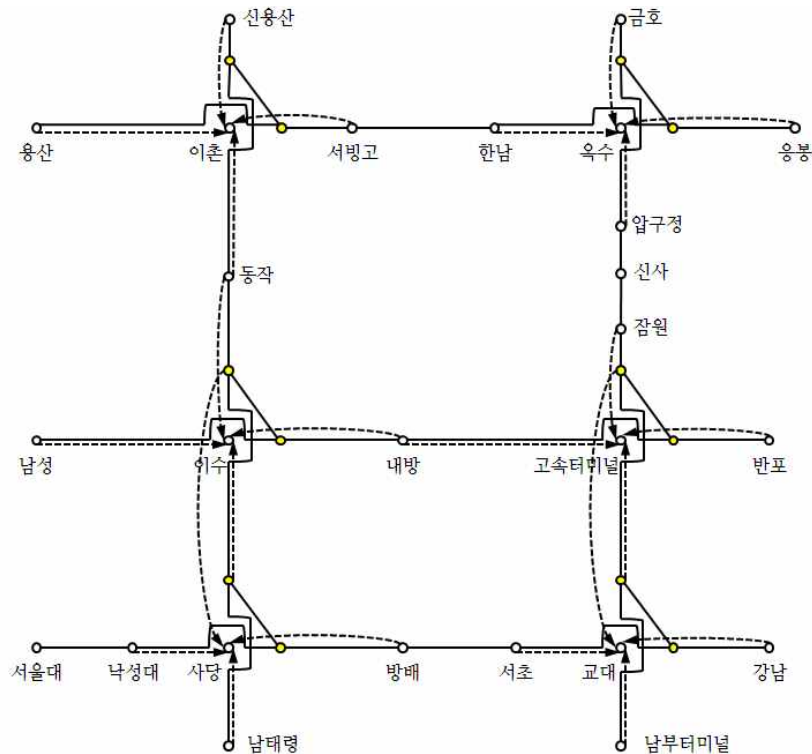
경로 A만 존재하는 결과가 도출된다. 한편, 고속터미널(3)을 도착지로 선택하여 분석하면 경로 B가 효율적 경로에 포함되기는 하지만, 경로 A가 고속터미널(7)→고속터미널(3) 환승링크 길이만큼을 총통행비용에 추가적으로 고려하는 문제가 생긴다. 물론 고속터미널(7)→고속터미널(3) 환승링크 길이를 0에 가까운 값을 사용하여 분석할 수는 있지만 이럴 경우 고속터미널을 제외한 다른 도착지에 대해 부정확한 결과를 도출할 수 있다. 즉, 고속터미널(7)→고속터미널(3) 환승링크 길이를 0에 가까운 값을 사용하면 하나의 출발지(서울대입구역)와 하나의 도착지(고속터미널)만을 정확하게 분석하고 고속터미널(7)→고속터미널(3) 환승링크를 통행하는 다른 도착지에 대해서는 부정확한 결과를 도출한다.

위에서 서술한 바와 같이 <그림 III-13>의 네트워크 표현 방식으로 는 도착지가 환승역인 경로를 분석하는데 적절하지 않다. 본 연구에서는 하나의 출발지에서 모든 도착지까지 환승통행을 포함한 모든 구간을 분석하기 위해 <그림 III-14>와 같은 네트워크 표현 방식을 제안한다. 네트워크에서 실선은 양방향으로 통행이 가능한 링크이고, 화살표식이 있는 점선은 화살표 방향으로만 통행이 가능한 아크(arc)를 의미한다.



<그림 III-14> Dial 알고리즘을 적용하기 위한 네트워크 표현 방식

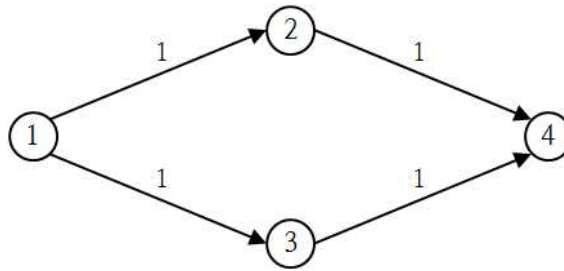
<그림 III-14>의 네트워크 표현 방식을 이용하여 다시 서울대입구역에서 고속터미널까지의 효율적 경로를 분석하면, 경로 A와 경로 B는 각각 내방역과 교대(3)역까지 실선의 링크를 이용해 통행하고 고속터미널까지는 점선의 아크를 이용해 통행을 종결하여 두 경로 모두 선택확률을 구할 수 있다. 이처럼 <그림 III-14>의 네트워크 표현 방식은 환승통행을 반영할 수 있으면서 환승역이 도착지인 구간을 분석하기가 용이하다. 게다가 별다른 수정없이 고속터미널 뿐만 아니라 다른 모든 도착역까지 한번에 통행배정할 수 있다. 서울대입구역이 출발지가 아닐 때에도 분석하기 위하여 <그림 III-14>의 네트워크 표현 방식을 보다 일반화하여 나타내면 <그림 III-15>와 같다. 하지만 <그림 III-15>의 일반화된 네트워크에서도 출발지에 따라 네트워크를 수정해야하는 문제가 있다. 요컨대 출발지가 환승역이라면, <그림 III-15>의 네트워크 표현방식에서 출발역 연결을 <그림 III-12>에서의 표현 방식으로 수정해야 한다.



<그림 III-15> 일반화된 네트워크 표현 방식

2) 출발지-도착지 별 링크이용확률

효율적 경로의 두 번째 정의를 이용한 Dial 알고리즘은 한 번의 수행으로 하나의 출발지에서 모든 도착지에 대해 통행을 배정한다. 이러한 특징은 계산적으로 효율적이기는 하지만, 특정 출발지-도착지에 대하여 링크(i, j)를 이용하는 통행량 또는 링크의 이용확률을 분석할 수 없다. 예를 들어, 링크의 길이가 모두 같은 <그림 III-16>의 예제 네트워크(10)에서 노드 ①을 출발지로 하고 각 도착지를 ②, ③, ④라고 했을 때, 각 링크의 이용확률은 도착지에 따라 달라진다. 이를테면, 링크(①, ②)의 이용확률은 도착지가 ②일 경우에는 1.0이고, 도착지가 ④일 경우에는 0.5가 된다. 그러나 효율적 경로의 두 번째 정의를 이용한 Dial 알고리즘은 단순히 노드에 연결된 링크들의 상대적 통행배정 확률을 구할 뿐 특정 도착지에 대한 링크의 이용확률을 구하지 않는다. 링크에 배정된 전체적인 통행량의 정보는 알 수 있지만 특정 도착지까지 어떤 링크 구간을 얼마만큼 이용하는지는 알 수 없다.



<그림 III-16> 예제 네트워크(10)

이러한 Dial 알고리즘의 취약점은 Van Vliet(1981)의 알고리즘을 통해 극복할 수 있다. Van Vliet(1981)은 최단경로와 통행거리가 Δp 만큼 차이 나는 경로 p 의 이용확률 $\Pr(p) = \exp(\theta^k \Delta p) / \sum_{p \in P_{o,d}^k} \exp(\theta^k \Delta p)$ 에서 $\exp(\theta^k \Delta p)$ 를 간단히 ϵ_p 로 표기하고 경로 p 의 유효성(effectiveness)이라

하였다. 그러면 경로 p 의 이용확률 $\Pr(p)$ 는 $\Pr(p) = \epsilon_p / \sum_{p \in P_{o,d}^k} \epsilon_p$ 로 다시 쓸 수 있다.

경로 p 를 $p_1: o \sim i$, $p_2: i \sim j$, $p_3: j \sim d$ 의 세 구간으로 나눌 수 있다. 면 Δp 는 $\Delta p = \Delta p_1 + \Delta p_2 + \Delta p_3$ 또는 $\Delta od = \Delta oi + \Delta ij + \Delta jd$ 이다. 여기서, $\Delta ij = p(j) - p(i) - t(i, j)$ 이다. 경로 p 의 이용확률은 도착지 d 에서 출발지 o 로 진행하며 각 구간(또는 링크)의 이용확률을 곱하여 구할 수 있다.

$$\begin{aligned}
 \Pr(p) &= \frac{w(j, d)}{W(d)} \times \frac{w(i, j)}{W(j)} \times \frac{w(o, i)}{W(i)} \\
 &= \frac{W(j) \exp(\theta \Delta jd)}{W(d)} \times \frac{W(i) \exp(\theta \Delta ij)}{W(j)} \times \frac{W(o) \exp(\theta \Delta oi)}{W(i)} \\
 &= \exp(\theta \Delta jd) \times \exp(\theta \Delta ij) \times \exp(\theta \Delta oi) \times W(o) / W(d) \\
 &= \exp(\theta \Delta od) / W(d) \\
 &= \epsilon_p / W(d) \\
 &\because W(o) = 1
 \end{aligned}$$

또한, $\Pr(p) = \epsilon_p / \sum_{p \in P_{o,d}^k} \epsilon_p$ 이므로 $W(d) = \sum_{p \in P_{o,d}^k} \epsilon_p$ 이다.

k 번째 출발지-도착지 쌍에서 링크 $e = (i, j)$ 를 이용할 확률 $\Pr[(i, j) | (o, d)^k]$ 은 링크 $e = (i, j)$ 를 통과하는 경로들의 이용확률을 모두 합한 값과 같다. 따라서,

$$\Pr[(i, j) | (o, d)^k] = \sum_{p \in (o, \dots, i, j, \dots, d)} \epsilon_p / \sum_{p \in P_{o,d}^k} \epsilon_p.$$

$e = (i, j)$ 를 통과하는 경로 $p \in (o, \dots, i, j, \dots, d)$ 의 유효성 ϵ_p 는

$$\epsilon_p = \exp(\theta \Delta p) = \exp(\theta \Delta p_1) \times \exp(\theta \Delta p_2) \times \exp(\theta \Delta p_3) = \epsilon_{p_1} \times \epsilon_{p_2} \times \epsilon_{p_3}$$

이므로,

$$\begin{aligned} \sum_{p \in (o, \dots, i, j, \dots, d)} \epsilon_p &= \sum_{p_1} \sum_{p_2} \sum_{p_3} \epsilon_{p_1} \times \epsilon_{p_2} \times \epsilon_{p_3} \\ &= \sum_{p_1} \epsilon_{p_1} \times \sum_{p_2} \epsilon_{p_2} \times \sum_{p_3} \epsilon_{p_3} \\ &= W(i) \times W(j | W(i) = 1) \times W(d | W(j) = 1). \end{aligned}$$

따라서,

$$\Pr[(i, j) | (o, d)^k] = W(i) \times W(j | W(i) = 1) \times W(d | W(j) = 1) / W(d).$$

$W(j | W(i) = 1)$ 는 $W(i) = 1$ 을 제외한 모든 노드 가중치를 0으로 초기화한 후에 노드 i 이후의 구간에 대하여 Dial 알고리즘의 전방진행 단계를 다시 수행하여 구한 $W(j)$ 값이고, 마찬가지로 $W(d | W(j) = 1)$ 는 $W(j) = 1$ 을 제외한 모든 노드 가중치를 0으로 초기화한 후에 노드 j 이후의 구간에 대하여 Dial 알고리즘의 전방진행 단계를 다시 수행하여 구한 $W(d)$ 값이다. $W(i)$ 와 $W(d)$ 는 $W(o) = 1$ 을 제외한 모든 노드 가중치를 0으로 초기화한 후에 구한 값이므로 각각 $W(i | W(o) = 1)$, $W(d | W(o) = 1)$ 로 다시 표기할 수 있다.

Van Vliet(1981)이 제안한 방법을 통해 전통적인 Dial 알고리즘으로 분석하기 어려운 $\Pr(e^k)$ 를 구할 수 있음을 알아보았다. $\Pr(e^k)$ 는 k 번째 출발지-도착지 간 통행에서 링크 e 가 얼마의 확률로 이용되는지를 의미하며 경로분석 시에 중요한 자료로 쓰인다. 또한, $\Pr(e^k)$ 는 θ 의 정산에서 필요한 정보이기도 하다.

4. 파라미터 θ 의 정산

통행자의 경로선택 패턴을 잘 묘사하는 θ 값을 찾는 일은 알고리즘의 유용성과 직결된다. Dial 알고리즘의 θ 값을 정산하는 방법으로 최대우도 추정법을 이용한 Robillard(1974)의 연구가 있다. Robillard(1974)는 각 경로의 선택확률을 해당 경로를 이용한 통행자수(관측통행량)만큼 곱하여 출발지(o)-도착지(d) 간 관측통행 우도함수 $L(\bullet)$ 를 나타내었다.

$$L(\{x_p\}, \theta) = \prod_{p \in P_{o,d}} \{\text{Pr}(p)\}^{x_p},$$

$$\text{Pr}(p) = \exp \theta c(p) / \sum_{p \in P_{o,d}} \exp \theta c(p).$$

여기서, x_p 는 경로 p 를 이용하는 통행량, $\text{Pr}(p)$ 는 경로 p 의 선택확률, $c(p)$ 는 경로 p 의 통행거리를 의미한다.

파라미터 θ 는 하나의 출발지와 하나의 도착지마다 정산할 수 있다. 이를 위하여 위 식을 k 번째 출발지(o)-도착지(d) 쌍에서의 경로 $p \in P_{o,d}^k$, k 번째 출발지-도착지 쌍에서 경로 p 를 이용하는 관측통행량 x_p^k , k 번째 출발지-도착지 쌍의 θ 값인 θ^k 로 고쳐서 다시 나타낼 수 있다.

$$L(\{x_p^k\}, \theta^k) = \prod_{p \in P_{o,d}^k} \{\text{Pr}(p)\}^{x_p^k},$$

우도함수 $L(\bullet)$ 를 최대화하는 θ^k 의 추정치 θ^{k*} 를 구하기 위하여 편의상 우도함수 $L(\bullet)$ 에 로그를 취한 로그우도함수를 이용한다.

$$LL(\theta^k) = \ln L(\{x_p^k\}, \theta^k) = \sum_{p \in P_{o,d}^k} x_p^k \ln \text{Pr}(p).$$

로그우도함수 $LL(\bullet)$ 를 θ^k 에 대하여 미분한 도함수를 $f(\theta^k)$ 라 하면,

$$f(\theta^k) = \frac{dLL(\theta^k)}{d\theta^k} = \sum_{p \in P_{o,d}^k} x_p^k \Pr(p)^{-1} \frac{d}{d\theta^k} \Pr(p),$$

$$\frac{d}{d\theta^k} \Pr(p) = \Pr(p) \left\{ c(p) - \sum_{p \in P_{o,d}^k} c(p) \Pr(p) \right\}.$$

최대우도추정치 θ^{k*} 는 $f(\theta^k) = 0$ 이 되는 θ^k 의 해이다.

$$\begin{aligned} f(\theta^k) &= \sum_{p \in P_{o,d}^k} x_p^k \left\{ c(p) - \sum_{p \in P_{o,d}^k} c(p) \Pr(p) \right\} \\ &= \sum_{p \in P_{o,d}^k} x_p^k c(p) - \sum_{p \in P_{o,d}^k} x_p^k \sum_{p \in P_{o,d}^k} c(p) \Pr(p) \\ &= \sum_{p \in P_{o,d}^k} x_p^k c(p) - T^k \sum_{p \in P_{o,d}^k} c(p) \Pr(p) = 0, \end{aligned}$$

여기서, T^k 는 k 번째 출발지-도착지 간의 총통행량이다.

위 식에서 $\sum_{p \in P_{o,d}^k} x_p^k c(p)$ 는 관측된 총통행거리를, $T^k \sum_{p \in P_{o,d}^k} c(p) \Pr(p)$ 는 추정된 총통행거리를 의미한다. 경로단위로 표현된 두 식은 링크단위의 식으로 나타낼 수 있다.

$$\begin{aligned} \sum_{p \in P_{o,d}^k} x_p^k c(p) &= \sum_{e \in E_{o,d}^k} x_e^k c(e), \\ \sum_{p \in P_{o,d}^k} c(p) \Pr(p) &= \sum_{e \in E_{o,d}^k} c(e) \Pr(e^k). \end{aligned}$$

$f(\theta^k)$ 를 링크단위의 수식으로 다시 쓰면,

$$f(\theta^k) = \sum_{e \in E_{o,d}^k} x_e^k c(e) - T^k \sum_{e \in E_{o,d}^k} c(e) \Pr(e^k) = 0$$

여기서, E 는 경로를 구성하는 링크 e 의 집합이고, $\Pr(e)$ 는 $\Pr(p)$ 와 마찬가지로 θ 를 파라미터로 하는 함수이다.

최대우도추정치 θ^k 는 관측된 총통행거리와 예측된 총통행거리가 같게 되는 값을 알 수 있다. 한편, θ 를 각각의 출발지-도착지 쌍마다 달리 하는게 아니라 모든 출발지-도착지 쌍에 대해 같은 값을 쓴다면, 즉, 모든 k 에 대하여 $\theta^k = \theta$ 이면 최대우도추정치 θ^* 는 아래의 식과 같이 $f(\theta) = 0$ 이 되는 θ 의 해이다.

$$\begin{aligned} f(\theta) &= \sum_k f(\theta^k) \\ &= \sum_k \left\{ \sum_{e \in E^k} x_e^k c(e) - T^k \sum_{e \in E^k} c(e) \Pr(e^k) \right\} \\ &= \sum_{e \in E} \sum_k x_e^k c(e) - \sum_k T^k \sum_{e \in E} c(e) \Pr(e^k) \\ &= \sum_{e \in E} T_e c(e) - \sum_{e \in E} \left\{ c(e) \sum_k T^k \Pr(e^k) \right\} \\ &= \sum_{e \in E} c(e) \left\{ T_e - \sum_k T^k \Pr(e^k) \right\} = 0. \end{aligned}$$

여기서, T_e 는 링크 e 에서 관측되는 총통행량이고 $\sum_k T^k \Pr(e^k)$ 는 모든 출발지-도착지 쌍에 대하여 링크 e 에 배정되는 총통행량이다.

Robillard(1974)는 $f(\theta_1) \times f(\theta_2) \leq 0$ 인 θ_1 와 θ_2 의 값 사이에 존재하는 $f(\theta) = 0$ 의 해 또는 근사값을 얻을 때까지 반복적으로 θ_1 과 θ_2 의 간격을 좁혀가는 방법을 사용하였다. 본 연구도 Robillard(1974)의 방법을 준용하여 θ 를 정산한다. 단, $\Pr(e^k)$ 는 Van Vliet(1981)의 결과로 계산한다.

IV. 알고리즘의 성과도 평가

1. 개요

본 장에서는 대중교통 통행배정 알고리즘의 성과도 평가를 수행한다. 서론에서부터 논의한 바와 같이, 대중교통 통행배정 알고리즘은 링크기반 결정적 알고리즘, 경로기반 확률적 알고리즘, 링크기반 확률적 알고리즘으로 분류하는데, 각 유형별 대표적인 알고리즘인 최적전략 알고리즘, 로짓기반 알고리즘, 본 연구를 통해 수정 및 보완된 Dial 알고리즘의 성과도 평가를 통해 세 알고리즘의 적합성을 살펴본다. 성과도 평가는 예제 네트워크와 실제 네트워크에서 이루어 진다.

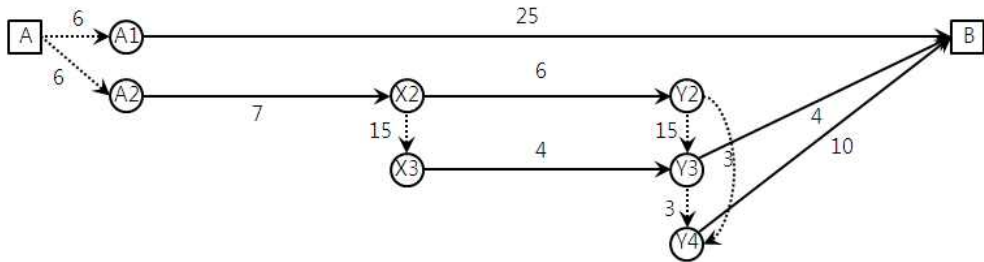
먼저, 다음의 2절에서는 가상적으로 설계된 경로특성과 가중치를 적용한 예제 네트워크를 기반으로 각 알고리즘을 비교, 분석한다. 예제 네트워크는 Spiess and Florian(1989)가 최적전략 알고리즘을 설명하기 위해 제시한 <그림 II-1>의 예제 네트워크(1)을 차용한다. 3절에서는 예제 네트워크 상에서 살펴본 각 알고리즘의 평가를 토대로 실제 수도권 도시철도 네트워크 중 일부 출발지-도착지 쌍을 선정하여 알고리즘의 성과도를 평가한다.

2. 예제 네트워크 기반 평가

1) 예제 네트워크 구축

예제 네트워크는 Spiess and Florian(1989)가 최적전략 알고리즘을 설명하기 위해 제시한 <그림 II-1>을 차용한다. 예제 네트워크(1)을 경로기반 확률적 알고리즘과 링크기반 확률적 알고리즘을 적용하기 위해 수정하여 나타내면 <그림 IV-1>의 예제 네트워크(11)과 같다. 실선 링

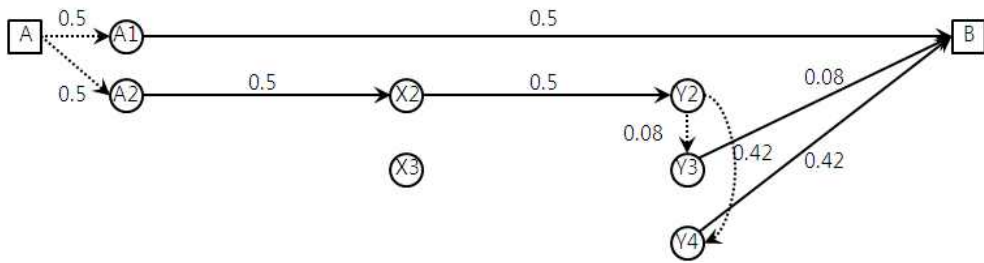
크는 차내통행을, 점선 링크는 대기 및 환승통행을 의미하며 각 링크 위의 숫자는 각 통행에서 소요되는 통행시간을 나타낸다. 대기시간은 배차 간격의 1/2을 사용하며, 차내시간 기준 대기시간 가중치 값에 따라 점선 링크의 저항은 달라진다. <그림 IV-1>에서는 대기시간 가중치를 1로 적용하였다.



<그림 IV-1> 예제 네트워크(11)

2) 최적전략 알고리즘 적용

먼저, 예제 네트워크(8)에 최적전략 알고리즘을 적용하여 링크의 이용확률을 계산해 보면 <그림 IV-2>와 같다. II장의 선행연구 고찰에서 살펴본 바와 같이, 각 링크(또는 노선)의 이용확률은 대기시간 가중치로 1이상의 어떤 값이 적용되어도 변하지 않는데, 그 이유는 예제 네트워크(8)에서 노선의 운행빈도가 고정되어 있기 때문이다. 달리 말하면, 대기시간 가중치 값에 따라 대기시간 비율이 큰 Line2, Line3, Line3로 구성된 경로와 상대적으로 대기시간 비율이 작은 Line1 경로의 총통행비용 차이가 달라지더라도 각 경로의 이용 확률은 운행빈도 비율에 따라 고정되어 나타난다. 이는 비합리적인 통행배정 결과이다. 다만 대기시간 비율의 변화로 특정 Line이 최적전략에서 제외될 경우 경로 집합과 이용확률이 달라질 수 있다.



<그림 IV-2> 링크이용확률(대기시간 가중치=1이상)

<표 IV-1> 출발지(A)-도착지(B) 간 경로 및 선택확률

구분	통과노드	총통행시간		경로선택 확률
		대기시간가중치=1	대기시간가중치=1.5	
경로1	A-A1-B	31.0	34.0	0.50
경로2	A-A2-X2-Y2-Y4-B	32.0	36.5	0.42
경로3	A-A2-X2-Y2-Y3-B	38.0	48.5	0.08

게다가 최적전략 알고리즘은 결정적 통행배정 알고리즘으로서 이용자의 통행시간에 대한 인지오차를 고려할 수 없다. 모든 이용자는 동일하게 최적경로로만 <표 IV-1>의 경로선택확률로 통행한다.

3) 경로기반 확률적 알고리즘 적용

다음으로 경로기반 확률적 알고리즘을 예제 네트워크에 적용한다. 우선 경로의 선택확률을 구하기 위해서는 경로선택에 영향을 미치는 속성변수들을 찾고, 효용함수를 추정해야 한다. 이 과정은 많은 시간과 비용을 필요로 하므로 선행연구 고찰(II장 3절 참고)에서 살펴본 효용함수를 차용하도록 한다. 선행연구에서 이경재(2004)와 Raveau et al.(2011)의 효용함수는 상향계단수, 하향계단수, 노선의 굴곡도 등 독특한 속성변수를 요구하고, 손상훈 외(2007)의 효용함수는 도보시간, 대기시간, 환승시

간을 속성변수로 사용하는데 환승에 소요되는 도보시간을 환승시간으로, 환승대기시간을 포함한 대기시간을 대기시간으로 구분하여 정의한다는 점에서 추가적인 조사가 필요하다. 또한 천승훈(2010)의 효용함수는 수단 별, 환승유무 별로 시장분할하여 각기 추정하기 때문에 경로 별로 효용함수를 달리 적용해야하는 번거로움이 있고, 유사경로의 정의를 통해 경로나열문제를 해결하고자 하였으나 유사경로 결정비용에 대한 타당한 근거가 제시되어 있지 않다. 이에 따라, 요구되는 변수의 종류 및 모형의 시장분할이 예제 네트워크 및 실제 네트워크에 적용이 용의한 양창화·손의영(2000)의 연구결과를 차용하도록 한다. 양창화·손의영(2000)의 경로 선택 효용함수 형태는 식 (2)와 같다.

$$U_i = \exp(V)$$

$$V_i = \beta_1 \times INTIME_i + \beta_2 \times TTIME_i + \beta_3 \times TNUMBER_i + \beta_4 \times TTIME_i \quad (2)$$

여기서, V_i 는 경로 i 의 효용함수이다. $INTIME_i$, $TTIME_i$, $TNUMBER_i$, $TTIME_i$ 는 각각 경로 i 의 차내시간, 환승시간, 환승횟수, 에스컬레이터 유무이고 β_1 , β_2 , β_3 , β_4 는 각 변수의 파라미터이다. 양창화·손의영(2000)은 파라미터를 SP자료와 RP자료를 통해서 추정하였는데 그 결과를 정리하면 <표 IV-2>와 같다.

<표 IV-2> 파라미터 추정결과

구분	파라미터(t 통계량)			
	차내시간(분)	환승시간(분)	환승횟수(회)	에스컬레이터유무 (무:0, 유:1)
SP자료	-0.184(-7.4)	-0.313(-7.3)	-1.881(-8.8)	0.368(2.7)
RP자료	-0.273(-14.3)	-0.281(-13.3)	-4.227(-16.0)	1.186(6.6)

양창화·손의영(2000)

효용함수 구축이 완료되면 출발지에서 도착지까지의 모든 경로를 나열한다. <그림 IV-1>의 예제 네트워크(8)에서 링크의 화살표 방향으로만 통행할 수 있다고 가정하면, 출발지-도착지 간 5개의 경로가 존재하며 <표 IV-3>과 같다.

<표 IV-3> 출발지(A)-도착지(B) 간 선택가능 경로집합 및 경로속성

구분	통과노드	차내시간 (분)	환승시간 ¹⁾ (분)	환승횟수 (회)	에스컬레이터 유무 ²⁾
경로1	A-A1-B	25	6	0	SP:0, RP:0
경로2	A-A2-X2-Y2-Y4-B	23	9	1	SP:1, RP:1
경로3	A-A2-X2-Y2-Y3-B	17	21	1	SP:1, RP:1
경로4	A-A2-X2-X3-Y3-B	15	21	1	SP:1, RP:1
경로5	A-A2-X2-X3-Y3-Y4-B	21	24	2	SP:1, RP:2

1) 양창화·손의영(2000)에서 환승시간은 환승을 위해 다른 노선으로 이동하는 시간과 차량대기시간을 합한 시간을 의미함

2) 양창화·손의영(2000)에서 에스컬레이터 유무는 SP모형에서는 유:1, 무:0으로, RP모형에서는 통행경로상에 에스컬레이터가 설치된 환승역 수로 입력함. 본 예제 네트워크는 모든 환승역에 에스컬레이터가 설비되어 있다고 가정함

경로의 선택확률은 양창화·손의영(2000)에서와 같이 로짓모형을 통해 구할 수 있으며, SP모형과 RP모형의 파라미터를 적용하여 구한 결과는 <표 IV-4>와 같다.

$$P_i = \frac{\exp(V_i)}{\sum_j \exp(V_j)}$$

여기서, P_i 는 경로 i 의 선택확률, j 는 대안경로의 수이다.

<표 IV-4> 경로기반 알고리즘의 경로선택확률

구분	경로선택확률	
	SP모형	RP모형
경로1	0.87250	0.94999
경로2	0.10857	0.03373
경로3	0.00766	0.00596
경로4	0.01106	0.01028
경로5	0.00022	0.00004

최적전략 알고리즘의 결과와 비교해 보았을 때 경로1의 선택확률은 증가하고 경로2, 경로3의 선택확률은 감소한 것을 알 수 있다. 또한 최적 경로에 포함되지 않는 경로4와 경로5를 선택하는 소수의 이용자를 설명한다. 이러한 차이는 최적전략 알고리즘이 노선의 배차간격 비율만으로 경로선택확률을 구하고, 최적경로에 전량 통행배정하는 결정적 알고리즘인 반면에, 경로기반 확률적 알고리즘은 이용자의 효용에 근거하여 경로 선택확률을 구하고, 최단경로가 아닌 경로에도 통행을 배정하는 확률적 알고리즘이라는 점에서 비롯한다.

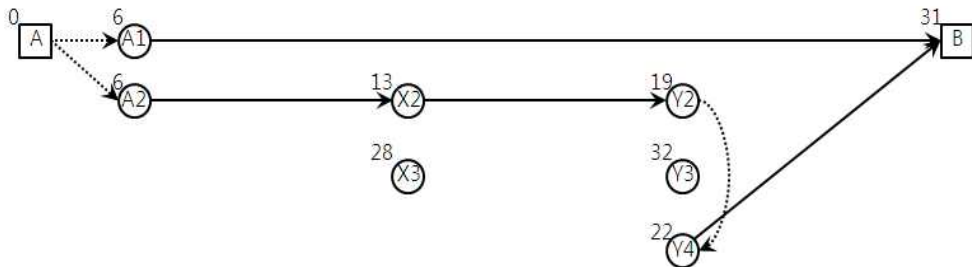
한편, 양창화·손의영(2000)의 SP모형과 RP모형으로 구한 경로의 선택확률은 최고 약 8%까지 차이가 난다. 선행연구의 효용함수를 정리한 <표 II-6>에서 연구마다 효용함수를 구성하는 속성변수와 추정된 파라미터의 크기가 다르다는 점을 감안한다면 효용함수에 따라 통행배정 결과는 민감하게 변할 수 있음을 알 수 있다.

또한 경로기반 알고리즘은 모든 경로를 나열해야 하기 때문에 계산적으로 많은 비용과 시간이 소요된다(본 예제 네트워크는 간단하고 방향도 한 방향으로 제한하여 이 문제가 부각되지는 않지만, 실제 네트워크에서는 상당히 많은 경로가 존재한다). 그 결과 환승이 잦고, 총통행시간이 비교적 커 이용될 가능성이 작은 경로도 선택대안 집합에 포함된다.

이로 인해, 이용되지 않을 가능성이 높은 경로에도 미약하지만 통행이 분배되어 오차가 발생할 수 있다.

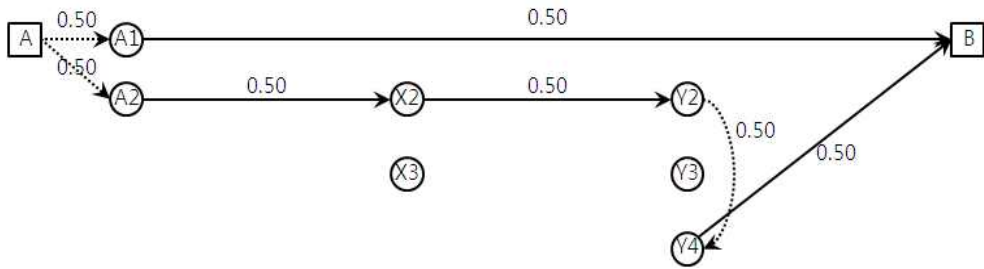
4) 링크기반 확률적 알고리즘 적용

마지막으로 링크기반 확률적 알고리즘인 Dial 알고리즘을 이용하여 링크(또는 경로)의 이용확률을 구하였다. Dial 알고리즘은 출발지로부터의 최단거리를 기준으로 전체 네트워크에서 효율적 경로를 판별한다. <그림 IV-3>은 출발지에서 각 역까지의 최단거리와 효율적 경로에 속하는 링크만을 나타내고 있다. 효율적 경로의 정의에 의하여 출발노드 A와 도착노드 B사이에는 총 2개의 경로만이 선택대안 집합에 포함된다.



<그림 IV-3> 최단경로거리와 효율적 경로(대기시간가중치=1)

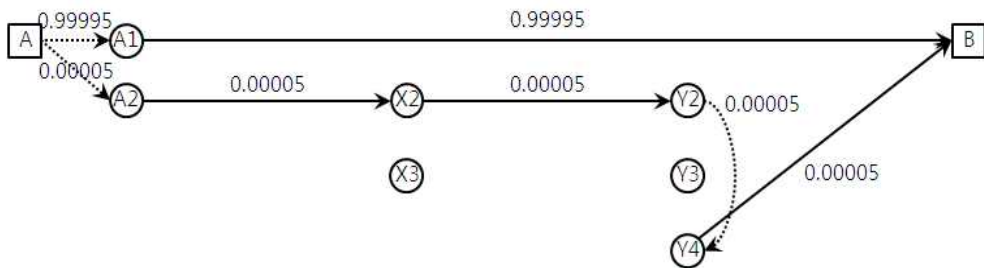
Dial 알고리즘은 인지오차 파라미터 θ 값을 적용하여 효율적 경로에 속한 링크의 선택확률을 구한다. θ 는 통행자의 인지오차 수준을 조절하기 위한 파라미터로서 값이 커질수록 통행자의 인지오차가 감소하여 최단경로를 선택할 확률이 커지고, 값이 작아질수록 통행자의 인지오차가 커져서 최단경로가 아닌 경로를 선택할 확률이 줄어든다. θ 값은 관측통행량 자료를 이용하여 정산한 값을 사용해야 하지만, 예제에서는 실제 관측통행량 정보가 없으므로 0, 1, 10을 예시적으로 적용한다. θ 값 별 링크의 이용확률을 구하면 각각 <그림 IV-4>, <그림 IV-5>, <그림 IV-6>과 같이 산출된다.



<그림 IV-4> $\theta=0$ 을 적용하여 구한 링크의 이용확률



<그림 IV-5> $\theta=1$ 을 적용하여 구한 링크의 이용확률



<그림 IV-6> $\theta=10$ 을 적용하여 구한 링크의 이용확률

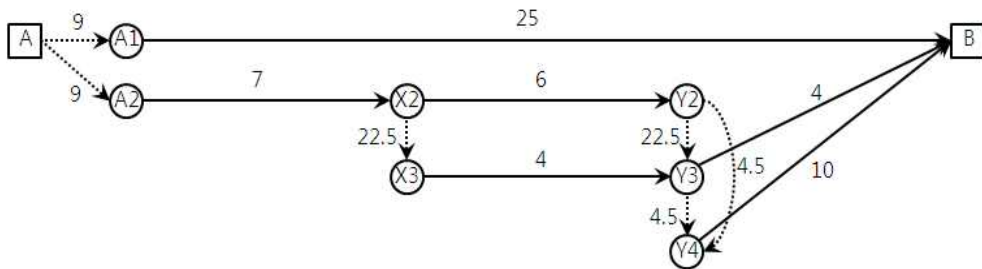
<그림 IV-4>, <그림 IV-5>, <그림 IV-6>에서 나타나는 결과를 통해 θ 값에 따라 링크의 이용확률이 달라지는 것을 실증적으로 확인할 수 있다. 링크의 이용확률을 경로의 이용확률로 정리하여 나타내면 <표 IV-5>와 같다. 경로1과 경로2는 총통행시간이 각각 31과 32로 1만큼 차이 나는데, θ 값이 0일 경우 통행자는 이 차이를 인지하지 못하고 두 경로를

동등한 확률로 선택하고, θ 값이 커질수록 통행자는 경로 간 통행시간 1 차이를 보다 잘 인지하게 되어 최단경로의 선택확률이 증가하는 것을 확인할 수 있다.

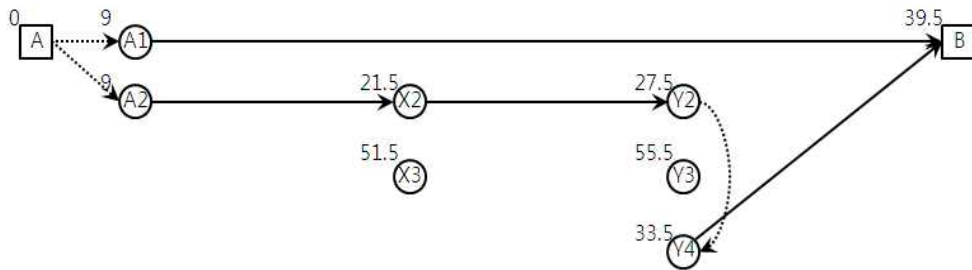
<표 IV-5> θ 값에 따른 경로선택 확률(대기시간 가중치=1)

구분	통과노드	총통행시간	경로선택확률		
			$\theta = 0$	$\theta = 1$	$\theta = 10$
경로1	A-A1-B	31.0	0.50000	0.73106	0.99995
경로2	A-A2-X2-Y2-Y4-B	32.0	0.50000	0.26894	0.00005

한편, 차내시간 기준 대기시간 가중치를 emme/3에서 기본값(default value)으로 사용하는 1.5로 적용하여 <그림 IV-1>의 예제 네트워크를 다시 나타내면 <그림 IV-7>의 예제 네트워크(12)와 같고 이에 대한 효율적 경로는 <그림 IV-8>과 같이 나타난다.



<그림 IV-7> 예제 네트워크(12)



<그림 IV-8> 최단경로거리와 효율적 경로(대기시간가중치=1.5)

대기시간 가중치로 1.5를 사용했을 때에도 θ 값을 0, 1, 10으로 적용하여 링크 및 경로의 이용확률을 산출하면 <표 IV-6>과 같다. θ 값이 증가함에 따라 최단경로의 선택확률이 증가하는 성향은 같으나, 대기시간 가중치가 1일 때보다 두 경로의 총통행시간 차이가 더 커지므로 같은 θ 값에서 최단경로의 선택확률이 더 크게 나타난다.

<표 IV-6> θ 값에 따른 경로선택 확률(대기시간 가중치=1.5)

구분	통과노드	총통행시간	경로선택확률		
			$\theta = 0$	$\theta = 1$	$\theta = 10$
경로1	A-A1-B	34.0	0.50000	0.92414	1.00000
경로2	A-A2-X2-Y2-Y4-B	36.5	0.50000	0.07586	0.00000

5) 소결론

지금까지 링크기반 결정적 알고리즘, 경로기반 확률적 알고리즘, 링크기반 확률적 알고리즘의 유형별 대표적인 알고리즘인 최적전략 알고리즘, 로짓기반 알고리즘, Dial 알고리즘을 예제 네트워크에 적용하여 경로의 선택확률을 구해 보았고, 그 과정에서 세 알고리즘의 특성을 살펴보았다. II장의 선행연구 고찰을 통해 살펴본 내용과 위의 과정을 통해 도

출한 결과를 종합하여 세 알고리즘을 평가하면 다음과 같다.

먼저, 링크기반 결정적 알고리즘은 통행자의 인지오차를 반영할 수 없어 이용자의 통행행태를 설명하는데 한계가 있다. 링크기반 결정적 알고리즘의 예로 살펴본 최적전략 알고리즘은 최적경로가 경우에 따라서 다중경로 집합으로 나타나고 각 경로는 이용확률을 갖지만, 이는 효용극대화이론을 기반으로 이용자의 인지오차를 고려하는 경로의 선택확률이 아니다. 경로의 선택확률은 운행빈도의 함수로서 분석가의 관점에서 정의하는 결정론적인 측면의 이용 ‘비율’이며 다른 서비스수준은 반영하지 못한다. 따라서 어떤 한 경로의 서비스수준이 달라지더라도 최적전략이 변화가 없다면 경로의 이용확률은 고정된 값으로 도출되어 비합리적이다.

경로기반 확률적 알고리즘은 출발지-도착지 간 모든 경로를 나열해야하기 때문에 대규모 네트워크에는 사실상 적용이 불가능하다. 또한, 모든 경로를 선택대안 집합에 포함하여 확률을 구하기 때문에 실제로 고려되지 않은 경로에도 통행이 배정된다. 또한 경로선택 시 판단근거가 되는 통행자의 효용함수를 추정하는 것은 많은 시간과 비용을 소요할 뿐만 아니라, 속성변수의 설정과 파라미터 추정결과에 따라 통행배정 결과가 민감하게 변할 수 있다.

반면에, 링크기반 확률적 알고리즘의 예로 살펴본 Dial 알고리즘은 위의 두 알고리즘이 갖는 문제점들을 극복한다. 우선, Dial 알고리즘은 확률적 통행배정 알고리즘으로서 모형상에서 파라미터 θ 를 이용하여 통행자의 인지오차를 고려할 수 있으며, 인지하는 통행시간의 차이에 따라 그 확률이 달라진다. 또한 Dial 알고리즘은 비합리적인 경로를 효율적 경로의 정의를 통해 사전에 제거하고, 링크와 노드 정보만을 이용하여 링크의 이용확률을 계산하는 링크기반 알고리즘이기 때문에 경로나열 문제가 발생하지 않는다. 경로기반 확률적 알고리즘은 다양한 속성변수를 이용해 효용함수를 구축해야한다는 점에 비해서, Dial 알고리즘은 단 하나의 파라미터 θ 만을 정산하면 된다는 장점도 있다.

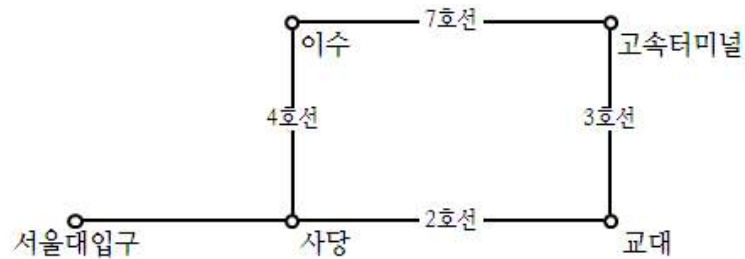
3. 실제 네트워크 기반 평가

1) 분석자료

앞서 2절에서는 가상의 예제 네트워크를 기반으로 한 인위적인 알고리즘의 비교를 진행했다면, 본 절에서는 실제 수도권 도시철도 네트워크에서 알고리즘의 성과도를 비교하고 각 알고리즘의 특성을 살펴보고자 한다. 알고리즘의 정산 및 검증을 위한 실제 통행량 데이터는 국가교통 DB센터에서 제공하고 한국철도공사 출처의 ‘월별 지하철 여객 수송실적 자료’ 중 2009년 4월, 2010년 4월, 2011년 4월 자료를 사용한다. 월별 지하철 여객 수송실적 자료는 출발역(O)-도착역(D) 간 월별 통행량 자료인데 본 연구에서는 일평균통행량으로 분석한다.

‘지하철 여객 수송실적 자료’는 환승역을 노선 별로 구분하는데, 예를 들어, 2호선과 4호선의 환승역인 사당역은 사당(2)와 사당(4)로 구분된다. 그런데 지하철 여객 수송실적 자료는 역 간 O/D자료이므로 이용자가 실제로 이용한 경로를 정확히 알 수 없다. 즉, 출발역-도착역 간 이용된 경로와 그 통행량을 알 수 없다. 본 연구는 이러한 한계점 속에서도 수도권 도시철도 네트워크 상에서 알고리즘을 비교하고 알고리즘의 특성을 살펴보기 위하여 경로 및 경로의 통행량을 합리적인 수준에서 판단할 수 있는 출발지-도착지 쌍을 선정한다. 이러한 지점으로 서울대입구-고속터미널 구간을 예로 들 수 있는데, <그림 IV-9>에서 나타나듯이, 서울대입구와 고속터미널 간에는 A경로(서울대입구-사당(2)-이수(4)-고속터미널(7))와 B경로(서울대입구-교대(2)-고속터미널(3))로 두 가지 경로만 존재한다고 판단하여도 직관에 반하지 않고, 각 경로(또는 경로를 구성하는 링크)를 이용한 실제 통행량은 ‘월별 지하철 여객 수송실적 자료’에서 각각 서울대입구-고속터미널(3), 서울대입구-고속터미널(7)의 수송량 정보를 통해 알 수 있다. 물론, A경로를 이용하더라도 고속터미널(7)에서 하차(alight)한 후 환승통로로 이동하여 고속터미널(3)에서

개찰(egress)하거나, 혹은 B경로를 이용한 후 고속터미널(7)에서 개찰할 수 있지만, 이러한 통행량은 무시할 수 있는 오차로 보아도 무방할 것이다.



<그림 IV-9> 서울대입구-고속터미널 간 네트워크

서울대입구-고속터미널 쌍을 포함하여, 분석지점으로 선정된 출발지-도착지 쌍은 <표 IV-7>과 같다. 앞서 설명한 바와 같이, 분석지점은 수도권 도시철도 네트워크(2009년 기준)와 지하철 여객 수송실적 자료로부터 경로 및 경로의 통행량을 파악할 수 있다고 판단되는 출발지-도착지 쌍이다.

<표 IV-7> 분석지점(O-D) 및 실제 이용된 경로 정보

O-D	선택대안경로 집합	차내거리 (km)	환승횟수 (회)	실제통행량(통행)		
				'09.4.	'10.4.	'11.4.
서울대입구 - 고속터미널	경로A(서울대입구-사당-이수-고속터미널)	7.0	2	26	42	35
	경로B(서울대입구-교대-고속터미널)	8.3	1	805	812	811
서울대입구 - 도곡	경로A(서울대입구-선릉-도곡)	11.6	1	61	61	59
	경로B(서울대입구-교대-도곡)	11.4	1	153	155	148
서울대입구 - 온수	경로A(서울대입구-신도림-온수)	14.6	1	25	20	23
	경로B(서울대입구-대림-온수)	13.9	1	38	55	52

<표 IV-7>에서 차내거리는 서울메트로, 서울도시철도공사, 한국철도공사 등의 역간거리 및 영업거리 정보를 활용하였다. 기타 지하철 평균 표정속도 및 평균환승시간, 노선 별 배차간격 자료는 <표 IV-8>, <표 IV-9>와 같으며 자료의 쓰임은 알고리즘마다 다르기 때문에 각 알고리즘의 적용에서 다시 설명하도록 한다.

<표 IV-8> 지하철 평균표정속도 및 평균환승시간

지하철 평균표정속도		지하철 간 평균환승시간
33.57(km/시) ¹⁾	559.5(m/분)	4.7분 ²⁾

1) 서울도시철도와 서울메트로의 총영업거리를 총소요시분으로 나누어 산정함

2) 국토해양부 보도자료(2010)

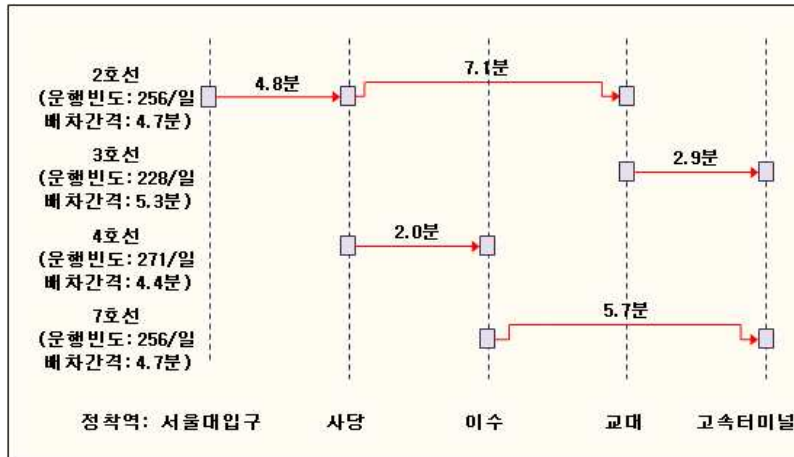
<표 IV-9> 지하철 노선 별 배차간격

구분	2009년		2010년		2011년	
	첨두	비첨두	첨두	비첨두	첨두	비첨두
1호선(경부)	5.7	7.2	5.7	7.2	6	7.1
2호선	2.5	6	2.5	6	2.5	6
3호선	3	6.5	3	6.5	3	6.5
4호선	2.5	5.5	2.5	5.5	2.5	5.5
5호선	2.5	6	3	6	2.5	6
6호선	4	8	5	8	4	8
7호선	2.5	6	2.5	6	2.5	6
8호선	4.5	8	5	8	4.5	8
분당선	5.5	9.2	4.1	7	4.1	7

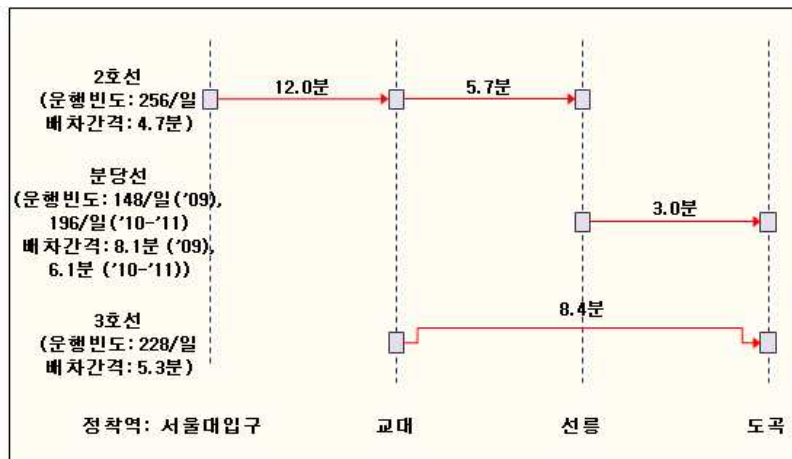
서울도시철도(2009), 서울도시철도(2010), 서울도시철도(2011)

2) 최적전략 알고리즘 적용

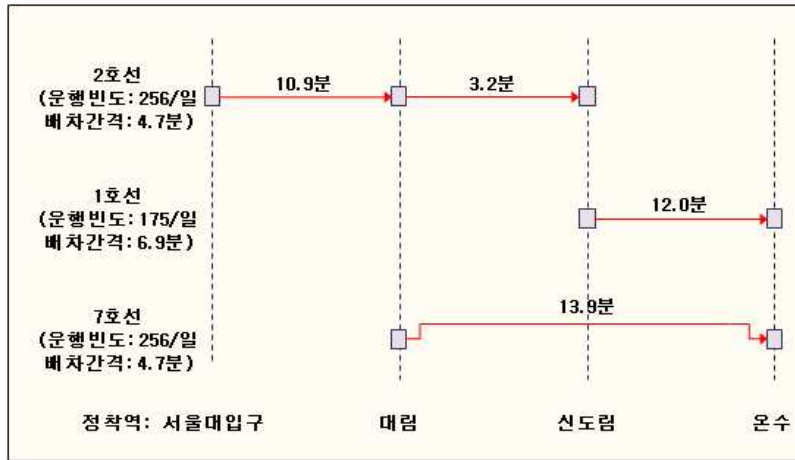
<표 IV-7>의 분석지점들에 대하여 선택가능한 노선(전략)을 나타내면 각각 <그림 IV-10>, <그림 IV-11>, <그림 IV-12>와 같다.



<그림 IV-10> 서울대입구-고속터미널 간 선택가능 노선(전략)



<그림 IV-11> 서울대입구-도곡 간 선택가능 노선(전략)



<그림 IV-12> 서울대입구-운수 간 선택가능 노선(전략)

<그림 IV-10>, <그림 IV-11>, <그림 IV-12>에서 각 노선의 차내통행시간은 역간거리를 지하철 평균표정속도로 나눈 값이고, 노선 별 운행빈도 및 배차간격은 <표 IV-9>에서 제시한 첨두 및 비첨두 시의 배차간격에 <표 IV-10>의 첨두 및 비첨두시간의 지속시간을 적용하여 구한 값이다. 노선 별 운행빈도와 배차간격은 <표 IV-11>에 정리하였다.

<표 IV-10> 수도권 지역의 1시간 통행량의 지속시간과 각 시간대별 집중률

구분	지속시간	집중률
첨두시간	4	8.78%
비첨두시간	16	3.84%
심야시간	4	-

한국개발연구원(2008)

<표 IV-11> 노선 별 하루(20시간)기준 운행빈도 및 배차간격

구분	2009년	2010년	2011년
1호선(경부)	175/일(6.9분)	175/일(6.9분)	175/일(6.9분)
2호선	256/일(4.7분)	256/일(4.7분)	256/일(4.7분)
3호선	228/일(5.3분)	228/일(5.3분)	228/일(5.3분)
4호선	271/일(4.4분)	271/일(4.4분)	271/일(4.4분)
7호선	256/일(4.7분)	256/일(4.7분)	256/일(4.7분)
분당선	148/일(8.1분)	196/일(6.1분)	196/일(6.1분)

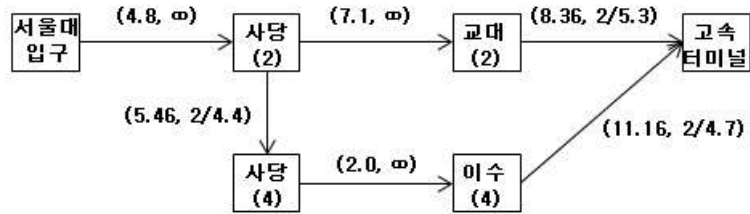
최적전략은 대기시간 등을 포함한 총예상통행시간을 최소화하는 전략의 집합이며, 총예상통행시간을 구하기 위해서는 승차시간(boarding time), 대기시간 인자(wait time factor), 대기시간 가중치(wait time weight), 승차시간 가중치(boarding time weight), 접근시간 가중치(auxiliary time weight)에 대한 정보가 필요하다. 여기서 승차시간은 환승을 포함한 모든 승차에 적용되는 페널티이고, 대기시간 인자는 배차간격으로 대기시간을 구할 때 사용되는 파라미터이다(INRO, 1998). 그런데 <표 IV-8>에서 제시한 지하철 간 평균환승시간은 환승에 소요되는 대기 및 이동시간을 모두 포함하는 값으로 최적전략 알고리즘의 파라미터에 직접적으로 적용하는 것은 부적절하다. 게다가 각 파라미터에 대해 명확히 구분하여 추정한 국내 연구는 찾아보기 힘들다. 따라서 각 파라미터 값은 Parveen et al.(2007)가 추정한 파라미터 값을 차용하는데, 이 값은 토론토 트랜짓 네트워크를 대상으로 실제 통행량과 EMME/2로 구한 배정통행량의 오차를 최소화하는 값이다.

<표 IV-12> 파라미터 추정치

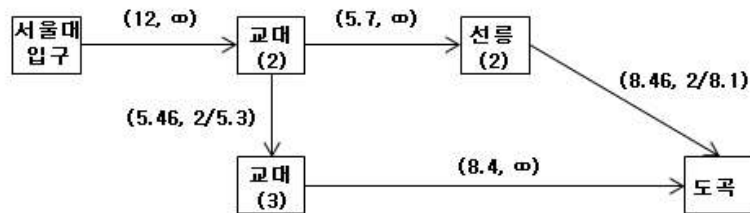
구분	파라미터 추정치
승차시간	2.6
대기시간 인자	0.5
대기시간 가중치	1.7
승차시간 가중치	2.1
접근시간 가중치	1

Parveen et al.(2007)

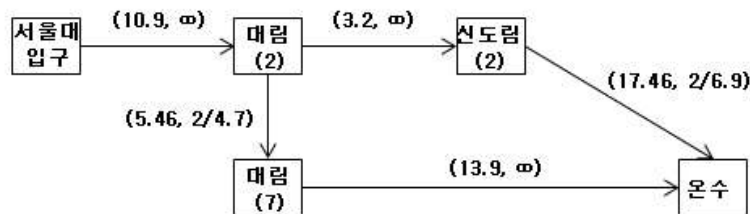
<그림 IV-10>, <그림 IV-11>, <그림 IV-12>에 Parveen et al.(2007)의 파라미터 추정결과를 적용하고, 최적전략을 탐색하기 용이한 간단한 네트워크(simplified transit network)로 다시 나타내면 각각 <그림 IV-13>, <그림 IV-14>, <그림 IV-15>와 같다.



<그림 IV-13> 서울대입구-고속터미널 간 네트워크



<그림 IV-14> 서울대입구-도곡 간 네트워크



<그림 IV-15> 서울대입구-온수 간 네트워크

링크는 화살표로 표시하고, 링크 상단에는 (링크통행저항 c_a , 운행빈도 f_a)를 표기하였다. 위의 네트워크에 표시된 운행빈도는 2009년 자료이며, 배차간격에 대기시간 인자 0.5를 곱한 값의 역수를 사용한다. 운행빈도가 ∞ 인 링크는 대기시간이 없음을 의미하고 환승이 이루어지지 않는 구간이다. 출발역인 서울대입구역에서는 승차시간과 대기시간을 고려

하지 않고, 환승이 포함된 링크에는 환승패널티 5.46분(승차시간 2.6 × 승차시간가중치 2.1)을 추가적으로 고려한다. 예를 들어, <그림 IV-15>에서 대림(2)-대림(7) 간 환승통행은 5.46분의 환승패널티와 대기시간 2.35분(=4.7/2)이 적용된다. 신도림(2)-온수 간에는 신도림(2)-온수 간 차내시간 12분에 신도림 역에서의 환승패널티 5.46분을 더한 총 17.46분을 통행저항으로 적용하고, 대기시간은 3.45분(=6.9/2)이다. 대기시간은 총예상통행시간 u_i 을 계산할 때 1.7만점의 가중치를 추가적으로 고려해준다.

위와 같이 네트워크 자료가 구축되면 최적전략을 탐색한다. 각 출발지-도착지 쌍에서 최적전략 탐색 결과는 각각 <표 IV-13>, <표 IV-14>, <표 IV-15>로 나타난다.

<표 IV-13> 서울대입구-고속터미널 간 최적전략($a \in \bar{A}$)

#	Node labels (u_i, f_i)						Link with Min $u_j + c_a$			
	서울대입구	사당(2)	사당(4)	교대(2)	이수(4)	고속터미널	$a(i, j)$	f_a	$u_j + c_a$	$a \in \bar{A}$
1	$\infty, 0$	$\infty, 0$	$\infty, 0$	$\infty, 0$	$\infty, 0$	0, 0	교대(2), 고속터미널	2/5.3	8.36	y
2	"	"	"	12.87, 2/5.3	"	"	이수(4), 고속터미널	2/4.7	11.16	y
3	"	"	"	"	15.16, 2/4.7	"	사당(4), 이수(4)	∞	17.16	y
4	"	"	17.16, ∞	"	"	"	사당(2), 교대(2)	∞	19.97	y
5	"	19.97, ∞	"	"	"	"	사당(2), 사당(4)	2/4.4	22.62	n
6	"	"	"	"	"	"	서울대입구, 사당(2)	∞	24.77	y
최종	24.77, ∞	19.97, ∞	17.16, ∞	12.87, 2/5.3	15.16, 2/4.7	0, 0	-	-	-	-

<표 IV-14> 서울대입구-도곡 간 최적전략($a \in \bar{A}$)

#	Node labels (u_i, f_i)					Link with Min $u_j + c_a$			
	서울대입구	교대(2)	교대(3)	선릉(2)	도곡	$a(i, j)$	f_a	$u_j + c_a$	$a \in \bar{A}$
1	$\infty, 0$	$\infty, 0$	$\infty, 0$	$\infty, 0$	$0, 0$	교대(3), 도곡	∞	8.4	y
2	"	"	8.4, ∞	"	"	선릉(2), 도곡	2/8.1	8.46	y
3	"	"	"	15.35, 2/8.1	"	교대(2), 교대(3)	2/5.3	13.86	y
4	"	18.37, 2/5.3	"	"	"	교대(2), 선릉(2)	∞	21.05	n
5	"	"	"	"	"	서울대입구, 교대(2)	∞	30.37	y
최종	30.37, ∞	18.37, 2/5.3	8.4, ∞	15.35, 2/8.1	$0, 0$	-	-	-	-

<표 IV-15> 서울대입구-온수 간 최적전략($a \in \bar{A}$)

#	Node labels (u_i, f_i)					Link with Min $u_j + c_a$			
	서울대입구	대림(2)	대림(7)	신도림(2)	온수	$a(i, j)$	f_a	$u_j + c_a$	$a \in \bar{A}$
1	$\infty, 0$	$\infty, 0$	$\infty, 0$	$\infty, 0$	$0, 0$	대림(7), 온수	∞	13.9	y
2	"	"	13.9, ∞	"	"	신도림(2), 온수	2/6.9	17.46	y
3	"	"	"	23.33, 2/6.9	"	대림(2), 대림(7)	2/4.7	19.36	y
4	"	23.36, 2/4.7	"	"	"	대림(2), 신도림(2)	∞	26.53	n
5	"	"	"	"	"	서울대입구, 대림(2)	∞	34.26	y
최종	34.26, ∞	23.36, 2/4.7	13.9, ∞	23.33, 2/6.9	$0, 0$	-	-	-	-

서울대입구-고속터미널, 서울대입구-도곡, 서울대입구-온수 간에서 각각 사당(2)-사당(4), 교대(2)-선릉(2), 대림(2)-신도림(2)는 최적전략에

포함되지 않는다. 따라서 출발지가 서울대입구이고 도착지가 고속터미널, 도곡, 온수일 때 모두 단일 최단경로에 전량 배정되는 결과가 도출된다. 2010년, 2011년의 배차간격은 분당선에 한하여 차이가 있는데 달라진 값을 적용하여도 2009년 배차간격을 적용하여 구한 결과와 동일하다.

<표 IV-16> 최적전략 알고리즘의 배정통행량과 오차율

도착역	경로	실제통행량			배정통행량(오차율)		
		'09.4.	'10.4.	'11.4.	'09.4.	'10.4.	'11.4.
고속 터미널	경로A	26	42	35	0(-1)	0(-1)	0(-1)
	경로B	805	812	811	831(0.03)	854(0.05)	846(0.04)
도곡	경로A	61	61	59	0(-1)	0(-1)	0(-1)
	경로B	153	155	148	214(0.4)	216(0.39)	207(0.4)
온수	경로A	25	20	23	0(-1)	0(-1)	0(-1)
	경로B	38	55	52	63(0.66)	75(0.36)	75(0.44)

위의 결과에서 제시하지 않았지만, 대기시간 가중치로 1을 적용하여도 그 결과는 같게 도출된다. 앞서 선행연구 고찰의 <그림 II-5>를 예로서 설명하였고 위의 결과에서도 확인할 수 있듯이, 최적전략 알고리즘은 통행자의 통행패턴에 대한 결정론적 가정으로 인해 단일 최단경로에 전량 통행배정할 수 있다. 또한, 설령 최적전략이 대안경로 A, B를 모두 포함하는 다중경로로 나타나더라도 서울대입구-고속터미널 쌍, 서울대입구-도곡 쌍에서는 오차가 더욱 커질 가능성이 있다. 최적전략 알고리즘의 경로선택확률은 운행빈도의 함수인데, <표 IV-11>에서 제시한 바와 같이 수도권 도시철도 네트워크에서 배차간격은 거의 비슷하여 경로의 선택확률이 동등하게 분배되기 때문이다. 이렇듯, 최적전략 알고리즘은 이용자 통행행태에 대해 결정론적 가정을 한다는 점과 경로의 선택확률이 운행빈도를 제외한 다른 서비스 수준은 고려하지 못한다는 점에서 이용자의 경로선택에 영향을 미치는 요소가 많은 복잡한 수도권 도시철도 네트워크 분석에 바람직하지 않은 것으로 판단된다.

3) 경로기반 확률적 알고리즘 적용

경로기반 확률적 알고리즘은 출발지-도착지를 연결하는 모든 경로를 나열한 후 추정된 경로의 효용함수를 이용해 경로의 선택확률을 구한다. 그런데 아무리 짧은 구간에서도 비합리적인 경로를 포함하여 무수히 많은 경로가 존재하여 사실상 모든 경로를 고려하는 것은 불가능하다. 본 연구에서는 경로나열 과정을 생략하고 <표 IV-7>에서 제시한 실제 이용된 경로만을 선택대안집합으로 고려한다. 또한 효용함수는 2절에서와 동일하게 양창화·손의영(2000)의 연구결과를 차용한다. 각 출발역-도착역의 대안경로에 대하여 양창화·손의영(2000)의 효용함수에서 요구하는 경로의 속성을 정리하면 <표 IV-17>과 같다.

<표 IV-17> 출발역-도착역 별 경로 및 경로속성

출발역	도착역	경로	차내시간(분) ¹⁾	환승시간(분) ²⁾	환승횟수	에스컬레이터 유무
서울대입구	고속터미널	경로A	12.5	9.4	2	RP:1 SP:1
		경로B	14.8	4.7	1	RP:1 SP:1
	도곡	경로A	20.7	4.7	1	RP:1 SP:1
		경로B	20.4	4.7	1	RP:1 SP:1
	온수	경로A	26.1	4.7	1	RP:1 SP:1
		경로B	24.8	4.7	1	RP:1 SP:1

1) <표 IV-7>의 차내거리를 지하철 평균표정속도로 나눠서 구함

2) <표 IV-8>에서 제시한 지하철 간 평균환승시간 4.7분을 환승횟수만큼 곱하여 산출함

3) 양창화·손의영(2000)에서 에스컬레이터 유무 변수의 입력방식은 SP모형과 RP모형이 서로 다름. SP모형은 에스컬레이터가 설비된 경우 1, 설비되지 않은 경우 0으로 입력하고 RP모형은 통행 경로상에서 에스컬레이터가 설치된 환승역 수로 입력함. 예를 들어, 서울대입구-고속터미널 쌍에서 경로A는 사당역과 충신대입구역에서 두 번 환승하는데 사당역에서 2호선→4호선 환승구간에는 에스컬레이터가 설비되어 있지 않음. 즉, 통행경로상의 두 환승역 중 에스컬레이터가 설비된 환승역은 한 곳임. 따라서 에스컬레이터 유무 변수값은 SP모형에서 1, RP모형에서 1임. 그 외의 출발지-도착지 쌍에서는 모든 경로가 환승역이 한 곳이고 에스컬레이터가 설비되어 있으므로 에스컬레이터 유무 변수 값은 SP, RP모형 모두 1임

출발지-도착지 간 선택가능한 대안경로집합을 구축하고 속성변수 자료의 수집이 완료되면, <표 IV-2>의 파라미터를 적용하여 경로의 선택 확률을 구한다. 그 결과는 <표 IV-18>과 같다. 또한, 경로의 선택확률로 각 연도별 배정통행량을 산출하고 오차율을 정리한 결과는 <표 IV-19>와 같다.

<표 IV-18> 경로기반 확률적 알고리즘의 경로선택확률

출발역	도착역	경로	경로선택확률	
			SP	R.P
서울대입구	고속터미널	경로A	0.051	0.007
		경로B	0.949	0.993
	도곡	경로A	0.484	0.476
		경로B	0.516	0.524
	온수	경로A	0.443	0.415
		경로B	0.557	0.585

<표 IV-19> 배정통행량과 오차율

도착역	경로	실제 통행량			배정통행량(오차율)					
					SP모형			RP모형		
		'09.4.	'10.4.	'11.4.	'09.4.	'10.4.	'11.4.	'09.4.	'10.4.	'11.4.
고속 터미널	경로A	26	42	35	42 (0.63)	44 (0.04)	43 (0.23)	6 (-0.77)	6 (-0.85)	6 (-0.82)
	경로B	805	812	811	789 (-0.02)	810 (0)	803 (-0.01)	825 (-0.02)	848 (0.04)	840 (0.04)
도곡	경로A	61	61	59	103 (0.7)	104 (0.71)	100 (0.7)	102 (0.67)	103 (0.68)	98 (0.67)
	경로B	153	155	148	111 (-0.28)	112 (-0.28)	107 (-0.28)	112 (-0.27)	113 (-0.27)	109 (-0.27)
온수	경로A	25	20	23	28 (0.12)	33 (0.66)	33 (0.44)	26 (0.05)	31 (0.56)	31 (0.35)
	경로B	38	55	52	35 (-0.08)	42 (-0.24)	42 (-0.2)	37 (-0.03)	44 (-0.2)	44 (-0.16)

위 <표 IV-18>의 결과에서 서울대입구-도곡, 서울대입구-온수 쌍에서는 대안경로 간 경로선택확률의 분배비율이 유사하게 나타난다. 두 출발지-도착지 쌍에서 대안경로의 환승횟수, 환승시간, 에스컬레이터 유무 속성은 같고, 단지 대안경로 간 차내시간 차이 0.3분, 1.3분만이 경로선택확률에 작용하기 때문이다. 효용함수가 요구하는 속성변수의 값이 서울대입구-도곡 또는 서울대입구-온수 쌍과 유사하게 나타나는 다른 출발지-도착지 쌍에 대해 분석하더라도 대안경로 간 선택확률의 분배비율은 서울대입구-도곡 또는 서울대입구-온수 쌍과 비슷하게 나타날 것이다. 즉, 효용함수의 속성변수 값이 유사한 대안경로를 갖는 출발지-도착지 쌍은 해당 출발지-도착지의 특징적인 통행패턴과 무관하게 대안경로 간 경로선택확률의 분배비율이 비슷하게 나타난다. 그러나 오차율은 그렇지 않다. 서울대입구-도곡, 서울대입구-온수 쌍에서 대안경로의 선택확률 분배비율이 유사하게 나타나지만 오차율은 서울대입구-도곡 쌍에서 크게 나타나는 것을 확인할 수 있다. 이는 서울대입구-도곡 쌍의 통행패턴은 효용함수에 포함된 속성변수만으로 설명하는 것이 불충분하거나, 혹은 속성변수의 실제 영향력이 추정된 파라미터의 크기와 다르기 때문임을 추론할 수 있다. 한편, 대안경로의 환승횟수가 다른 서울대입구-고속터미널 쌍에서는 SP모형과 RP모형의 경로선택확률 차이가 비교적 크게 나타나는데 이는 환승횟수 변수의 파라미터가 2배 이상 차이 나기 때문이다.

지금까지의 논의로부터 경로기반 확률적 알고리즘의 효용함수와 관련된 한계점을 알 수 있다. 결론적으로, 이용자의 경로선택에 영향을 미치는 속성변수의 종류와 그 영향력은 경로, 지역, 이용자, 혹은 기타(기상상태, 통행시각 등) 특성 등 다양한 요인에 의해 달라질 수 있음에도 불구하고, 광범위한 범위에서 추정한 효용함수 또는 다른 연구의 효용함수를 빌려 적용함으로써 오차가 발생할 수 있다. 물론, 통행특성이 유사한 출발지-도착지 별로 효용함수를 추정하면 이러한 한계점을 해결할 수 있지만, 효용함수가 유의하기 위해서는 각 출발지-도착지 별로 이용

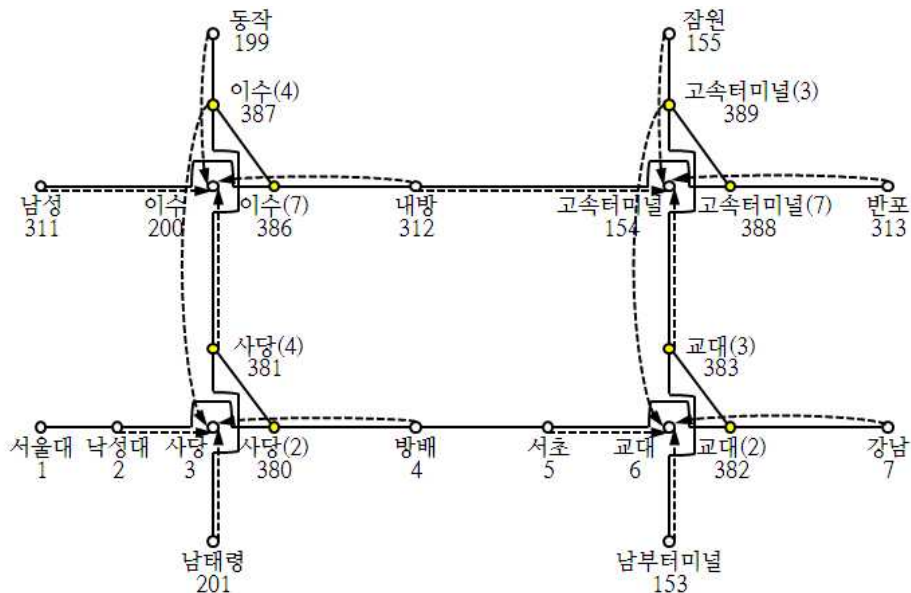
자 통행행태에 대한 많은 수의 표본을 확보해야만 한다.

한편, 비합리적인 경로를 포함한 모든 경로를 나열할 수 없다는 점 때문에 본 절에서 구체적으로 논의하지는 못했지만 경로기반 알고리즘은 분석단위가 경로라는 점에서 내제적으로 경로나열문제를 갖는다. 출발지-도착지 간 무수히 많은 경로 중에서 실제 이용될 가능성이 높은 경로만으로 선택대안집합을 구축하는 합리적인 방법론에 대한 논의가 필요하다.

4) 링크기반 확률적 알고리즘 적용

가) 네트워크 구축

Dial 알고리즘을 적용하기 위해서는 링크와 노드 정보 뿐만 아니라 노드 i 에서 시작하는 모든 링크의 집합 I_i , 노드 i 에서 종결하는 모든 링크의 집합 F_i , 출발역에서 노드 i 까지의 최단경로거리 $p(i)$ 등의 네트워크 정보가 필요하다. 우선 III장 3절에서 소개한 네트워크의 표현 방식을 적용하여 노드번호를 설정하고 노드 i 에서 시작하는 모든 링크의 집합 I_i 와 노드 i 에서 종결하는 모든 링크의 집합 F_i 데이터를 구축한다. 각 노드 i 에 연결된 초기노드 또는 최종노드의 개수(즉, 링크의 개수) 정보와 이들의 노드번호 정보를 함께 구축하고, 이 정보를 알고리즘 상에서 링크정보로 입력받아 링크의 집합인 I_i , F_i 데이터를 효율적으로 계산한다. <그림 IV-16>은 노드번호를 설정한 예시이며, <표 IV-20>과 <표 IV-21>은 해당 네트워크에 구축된 I_i , F_i 데이터 내역이다.



<그림 IV-16> 노드번호 설정(예시)

<표 IV-20> I_i 데이터 구축 예시(노드 i 에 연결된 최종노드의 개수와 노드번호)

노드번호(i)	개수	최종노드			
1	2	2	43(봉천)	-	-
2	3	1	3	380	-
3	0	-	-	-	-
4	3	3	5	380	-
5	3	4	6	382	-
6	0	-	-	-	-
7	3	6	8(역삼)	382	-
153	3	6	152(양재)	383	-
154	0	-	-	-	-
155	3	154	156(신사)	389	-
199	4	198(이촌)	200	238(이촌(4))	387
200	0	-	-	-	-
201	3	3	202(선바위)	381	-
380	3	2	4	381	-
381	4	200	201	380	387
382	3	5	7	383	-
383	4	153	154	382	389
386	3	311	312	387	-
387	4	3	199	381	386
388	3	312	313	389	-
389	4	6	155	383	388

<표 IV-21> F_i 데이터 구축 예시(노드 i 에 연결된 초기노드의 개수와 노드번호)

노드번호(i)	개수	초기노드			
1	2	2	43(봉천)		
2	2	1	380		
3	4	2	4	201	387
4	2	5	380		
5	2	4	382		
6	4	5	7	153	389
7	2	8(역삼)	382		
153	2	152(양재)	383		
154	4	155	312	313	383
155	2	156(신사)	389		
199	2	238(이촌(4))	387		
200	4	199	311	312	381
201	2	202(선바위)	381		
380	3	2	4	381	
381	3	201	380	387	
382	3	5	7	383	
383	3	153	382	389	
386	3	311	312	387	
387	3	199	381	386	
388	3	312	313	389	
389	3	155	383	388	

차내통행링크의 저항은 서울메트로, 서울도시철도공사, 인천메트로, 한국철도공사 등의 역간거리 및 영업거리 정보를 활용하고 환승링크의 저항은 차내시간 기준 환승통행저항을 적용한다. 차내시간 기준 환승통행저항은 평균환승시간을 거리기준으로 환산한 후 문헌에서 나타나는 차내시간 기준 환승가중치의 평균을 곱하여 계산한다.

$$\text{차내시간기준 평균환승통행저항} = \text{평균환승통행저항} \times \text{평균환승가중치}$$

<표 IV-22>의 결과에서 나타나듯이, 차내시간 기준 환승통행저항의 평균은 3.4km이다.

<표 IV-22> 차내시간 기준 환승통행저항(거리기준)

지하철 평균표정속도 (=a)		지하철 간 평균환승시간 (=b)	평균환승저항 (거리기준) ¹⁾ (=a*b=c)	차내시간 기준 환승통행저항 (거리기준)	
33.57 (km/시) ²⁾	559.5 (m/분)	4.7분 ³⁾	2.6(km)	환승가중치: 1.18 (신성일(2010))	3.1(km)
				환승가중치: 1.37 (손상훈 외(2007))	3.6(km)
				환승가중치: 1.03	2.7(km)
				환승가중치: 1.70 (양창화·손의영(2000)) ⁴⁾	4.4(km)
				평균 환승가중치: 1.32	3.4(km)

1) 통상 총통행시간은 ‘역간거리/지하철평균속도+환승시간×환승가중치’로 구하는데, 본 연구에서는 통행저항을 역간거리를 사용하고 지금까지의 논의가 통행거리로 이루어졌음을 감안하여, 환승통행도 거리를 기준으로 계산함

2) 서울도시철도와 서울메트로의 총영업거리와 총소요시분을 이용하여 구함

3) 국토해양부 보도자료(2010)

4) RP와 SP자료를 이용하여 각각 가중치를 구함

나) 최단경로 및 최단거리 $p(i)$ 산정

Dial 알고리즘을 적용하기 위해 추가적으로 필요한 정보는 출발역에서 각 노드 i 까지의 최단거리 $p(i)$ 이다. 최단거리 $p(i)$ 는 Floyd(1962)의 최단경로 알고리즘을 통해 구하였다.

Step 1. (초기화: 링크길이 및 이전노드(backnode) 입력)

모든 노드 o, i 에 대해서 다음을 계산한다.

$$q(o, i) \leftarrow o$$

$$o \text{와 } i \text{가 연결된 링크일 경우 } p(o, i) \leftarrow p(o, i)$$

$$\text{그렇지 않을 경우 } p(o, i) \leftarrow \infty$$

Step 2. (최단경로 및 최단거리 계산)

서로 다른 모든 노드 o, i, j 에 대하여 다음을 계산한다.

$$p(o, j) + p(j, i) < p(o, i) \text{면 } p(o, i) \leftarrow p(o, j) + p(j, i), q(o, i) \leftarrow q(j, i)$$

여기서, $p(o, i)$ 는 어떤 출발역 o 에서 각 노드 i 까지의 최단경로 거리 행렬을, $q(o, i)$ 는 $o-i$ 간 최단경로에서 노드 i 의 이전노드 행렬을 의미한다.

Van Vliet(1978)은 Floyd(1962)의 알고리즘을 적용하기 위해서는 하나의 노드에 세 개 이하의 링크가 연결되어 있어야 하며, 이러한 점 때문에 복잡한 교통망에서는 효율적인 알고리즘이 아닐 수 있다고 하였다. 실제로 Floyd의 알고리즘을 수도권 도시철도 네트워크에 한 번만 수행했을 경우 연결된 링크가 많은 노드에서 오차가 발생하였다. 그러나 Floyd의 알고리즘을 2번 이상 반복하여 수행하면 이러한 문제점이 해결되었다.

본 연구에서 최단거리를 구하는 이유는 단순히 Dial 알고리즘의 입력자료를 구하기 위함이므로 여러 최단경로 알고리즘 간 비교우위에 대한 논의는 생략하도록 한다. 다만, Floyd의 알고리즘은 프로그래밍이 용이하고, 요구하는 데이터가 간소하여 본 연구에 무리없이 적용할 수 있으며, 그 결과도 연구대상 네트워크에서 문제가 없는 것을 확인하였다.

다음의 <표 IV-23>, <표 IV-25>는 알고리즘의 초기 입력값이고 <표 IV-24>, <표 IV-26>은 Floyd의 최단경로 알고리즘으로 구한 최단거리와 최단경로 이전노드(backnode) 결과값(일부)이다. <표 IV-23>의 초기값 중에서 300은 연결되지 않은 링크의 통행거리로 무한히 큰 값을 의미한다.

<표 IV-23> 출발역-도착역 간 최단거리 초기값(링크길이)

O/D	1	2	3	4	5	6	7	8	9	10
1	300	1	300	300	300	300	300	300	300	300
2	1	300	1.7	300	300	300	300	300	300	300
3	300	1.7	300	1.6	300	300	300	300	300	300
4	300	300	1.6	300	1.7	300	300	300	300	300
5	300	300	300	1.7	300	0.7	300	300	300	300
6	300	300	300	300	0.7	300	1.2	300	300	300
7	300	300	300	300	300	1.2	300	0.8	300	300
8	300	300	300	300	300	300	0.8	300	1.2	300
9	300	300	300	300	300	300	300	1.2	300	1.3
10	300	300	300	300	300	300	300	300	1.3	300

<표 IV-24> 출발역-도착역 간 최단거리 결과값(1)

O/D	1	2	3	4	5	6	7	8	9	10
1	300	1	2.7	4.3	6	6.7	7.9	8.7	9.9	11.2
2	1	300	1.7	3.3	5	5.7	6.9	7.7	8.9	10.2
3	2.7	1.7	300	1.6	3.3	4	5.2	6	7.2	8.5
4	4.3	3.3	1.6	300	1.7	2.4	3.6	4.4	5.6	6.9
5	6	5	3.3	1.7	300	0.7	1.9	2.7	3.9	5.2
6	6.7	5.7	4	2.4	0.7	300	1.2	2	3.2	4.5
7	7.9	6.9	5.2	3.6	1.9	1.2	300	0.8	2	3.3
8	8.7	7.7	6	4.4	2.7	2	0.8	300	1.2	2.5
9	9.9	8.9	7.2	5.6	3.9	3.2	2	1.2	300	1.3
10	11.2	10.2	8.5	6.9	5.2	4.5	3.3	2.5	1.3	300

<표 IV-25> 최단경로 이전노드 초기값

O/D	1	2	3	4	5	6	7	8	9	10
1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9	9	9	9
10	10	10	10	10	10	10	10	10	10	10

<표 IV-26> 최단경로 이전노드 결과값

O/D	1	2	3	4	5	6	7	8	9	10
1	1	1	2	3	4	5	6	7	8	9
2	2	2	2	3	4	5	6	7	8	9
3	2	3	3	3	4	5	6	7	8	9
4	2	3	4	4	4	5	6	7	8	9
5	2	3	4	5	5	5	6	7	8	9
6	2	3	4	5	6	6	6	7	8	9
7	2	3	4	5	6	7	7	7	8	9
8	2	3	4	5	6	7	8	8	8	9
9	2	3	4	5	6	7	8	9	9	9
10	2	3	4	5	6	7	8	9	10	10

<표 IV-26>의 결과는 Dial 알고리즘을 적용할 때 실질적으로 필요한 데이터는 아니다. 그러나 네트워크의 구축(연결)이 제대로 되었는지 또는 최단경로가 정확한지를 검토할 때 용이하게 사용할 수 있다. 한편, Dial 알고리즘에서 실질적으로 필요한 최단거리 결과인 <표 IV-24>에서 출발지와 도착지가 같은 경로거리 즉, $p(o, i)$ 에서 $i=o$ 일 때도 역시 최단거리가 300으로 나타나는데 이 값을 0으로 변경해야 한다. 이 것은 위의 알고리즘 과정을 거친 이후에 다음의 명령을 수행함으로써 간단하게 처리할 수 있다.

$$p(o, i) = \begin{cases} p(o, i) & \text{if } i \neq o \\ 0 & \text{otherwise} \end{cases}$$

그러면 <표 IV-24>의 최단거리 $p(o, i)$ 의 결과는 최종적으로 <표 IV-27>과 같다.

<표 IV-27> 출발역-도착역 간 최단거리 결과값(2)

O/D	1	2	3	4	5	6	7	8	9	10
1	0	1	2.7	4.3	6	6.7	7.9	8.7	9.9	11.2
2	1	0	1.7	3.3	5	5.7	6.9	7.7	8.9	10.2
3	2.7	1.7	0	1.6	3.3	4	5.2	6	7.2	8.5
4	4.3	3.3	1.6	0	1.7	2.4	3.6	4.4	5.6	6.9
5	6	5	3.3	1.7	0	0.7	1.9	2.7	3.9	5.2
6	6.7	5.7	4	2.4	0.7	0	1.2	2	3.2	4.5
7	7.9	6.9	5.2	3.6	1.9	1.2	0	0.8	2	3.3
8	8.7	7.7	6	4.4	2.7	2	0.8	0	1.2	2.5
9	9.9	8.9	7.2	5.6	3.9	3.2	2	1.2	0	1.3
10	11.2	10.2	8.5	6.9	5.2	4.5	3.3	2.5	1.3	0

다) 수정된 Dial 알고리즘 구축

네트워크 구축 및 최단거리 산정 과정이 완료되면 Dial 알고리즘을 적용하여 통행배정할 수 있다. 그런데 앞서 III장 3절에서 살펴보았듯이, Dial 알고리즘을 통해 배정되는 링크의 통행량은 하나의 출발지에서 모든 도착지로 가는 전체 통행량으로서 전반적인 통행량의 흐름을 분석하는데는 용이하지만, 하나의 출발지와 하나의 도착지 쌍에 대해 이용된 경로와 그 이용확률을 분석할 수 없다. 본 연구에서는 전통적 Dial 알고리즘이 갖는 이러한 한계점을 해결하고자 Dial 알고리즘의 전방진행단계와 Van Vliet(1981)이 제시한 링크이용확률의 산출방법을 결합한 수정된 형태의 Dial 알고리즘을 구축하고 이를 통해서 출발지-도착지 별 이용된 경로 및 이용확률, 배정통행량 등을 분석한다. 수정된 알고리즘에서 사용되는 변수의 표기법은 다음과 같다.

총노드개수: 네트워크를 구성하는 총노드수

총통행량(o, d): 출발지 o 에서 도착지 d 로 통행하는 총통행량

링크(i, j): 초기노드 i 와 최종노드 j 를 연결하는 링크

링크길이(i, j): 링크(i, j)의 통행거리

초기노드(i, n_1): 노드 i 에서 종결하는 링크들을 구성하는 초기노드집합 중에서 n_1 번째 초기노드

최종노드(i, n_2): 노드 i 에서 시작하는 링크들을 구성하는 최종노드집합 중에서 n_2 번째 최종노드

초기노드수(i): 노드 i 에서 종결하는 링크들을 구성하는 초기노드의 개수

최종노드수(i): 노드 i 에서 시작하는 링크들을 구성하는 최종노드의 개수

$l(i, j)$: 링크(i, j)의 우도(likelihood)

최단거리(o, i): 출발지 o 에서 노드 i 까지의 최단거리

$W(o, i)$: 출발지가 o 일 때 노드 i 의 가중치

$w(i, j)$: 링크(i, j)의 가중치

탐색노드(n_3): 출발지 o 에서 n_3 번째로 먼 노드

$NW(o, i, j)$: 출발지가 o 일 때 노드 i 에서 가중치를 초기화 즉,
 $W(i)=1$ 을 제외한 모든 노드가중치는 0으로 하여 새로
구한 노드 j 의 가중치

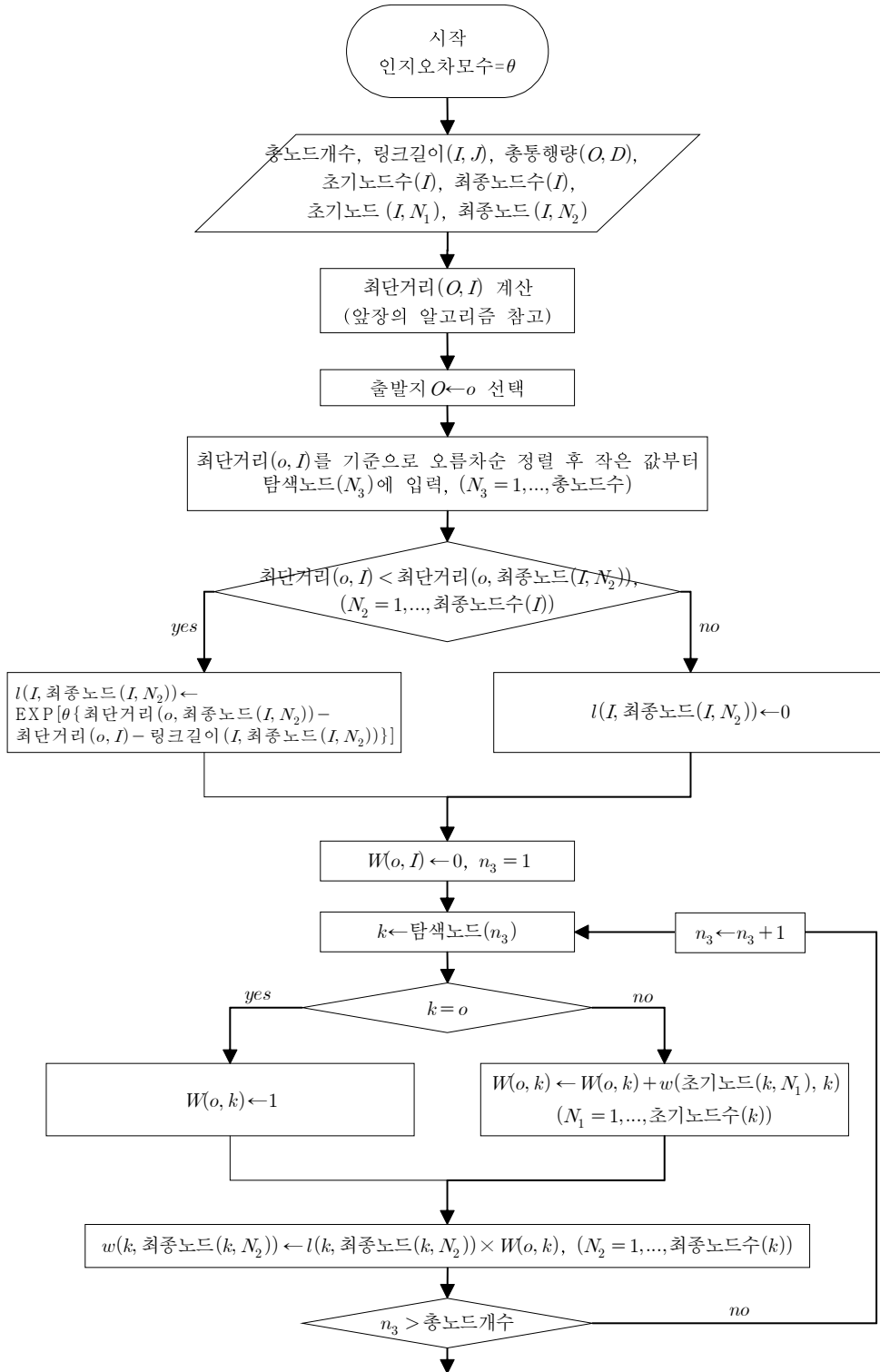
$nw(i, j)$: 초기화 후 새로 구한 링크(i, j)의 가중치

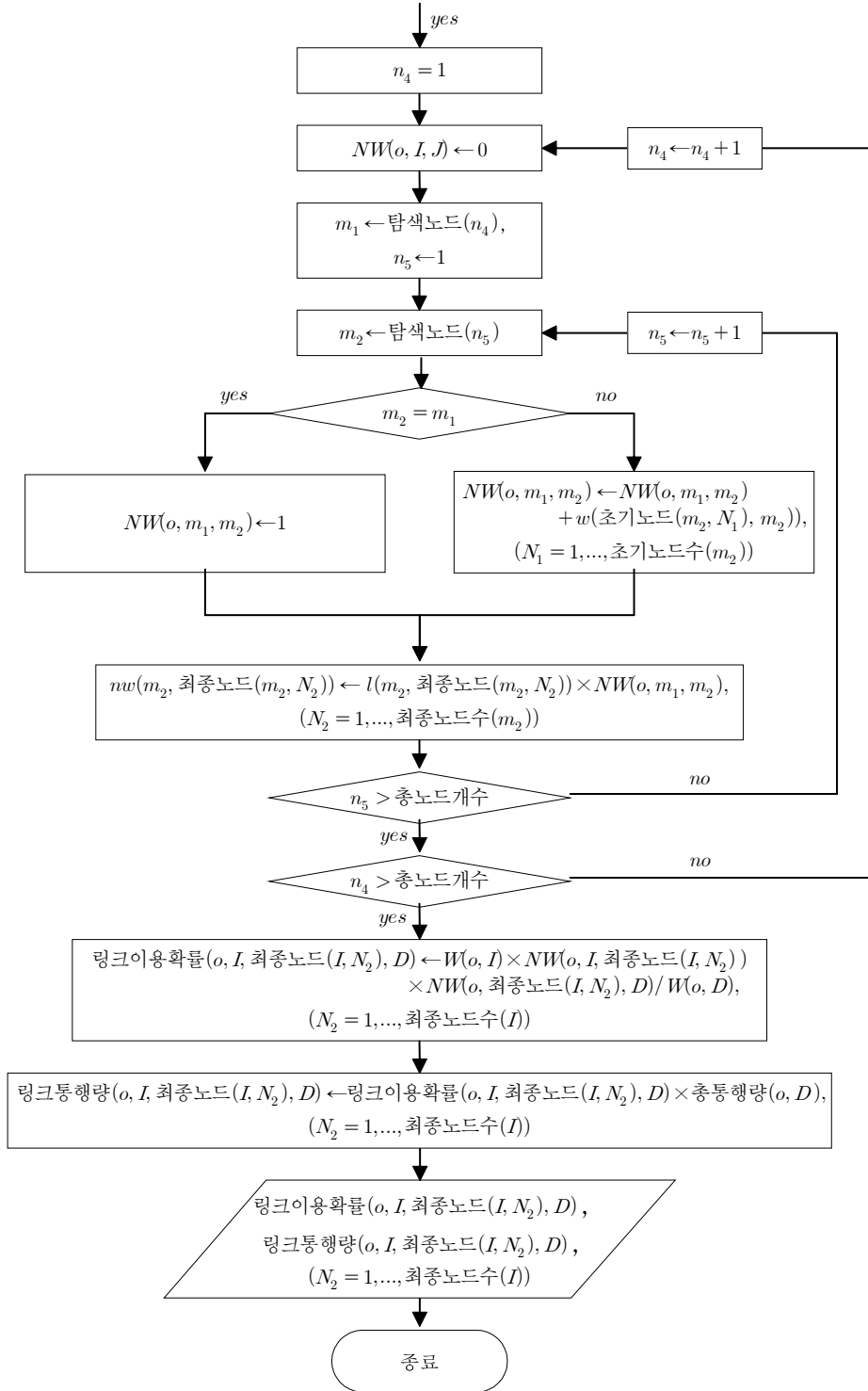
링크이용확률(o, i, j, d): 출발지가 o 이고 도착지가 d 일 때 링크(i, j)를
이용할 확률

링크통행량(o, i, j, d): 출발지가 o 이고 도착지가 d 일 때 링크(i, j)를
이용하는 통행량

위의 변수 중에서 알파벳으로 표기된 변수는 대문자로 표기할 때 해당 변수의 집합을 의미한다. 이를테면, I 또는 J 는 네트워크를 구축하고 있는 모든 노드 i 또는 j 의 집합을, O 와 D 는 출발지와 도착지 집합을, 링크길이(I, J)는 모든 링크에 대한 통행거리 정보를, 링크이용확률(o, i, j, D)는 출발지가 o 일 때 도착지 집합에 포함된 각각의 도착지 d 에 대해서 구한 링크(i, j)의 이용확률 집합을 의미한다. 단, 본래 대문자로 표기한 노드 가중치 $W(\cdot)$ 와 $NW(\cdot)$ 는 예외이다. 한편, 최단거리(o, o)=0, 탐색노드(1)= o , $W(o, i)=NW(o, o, i)$ 임은 쉽게 알 수 있다.

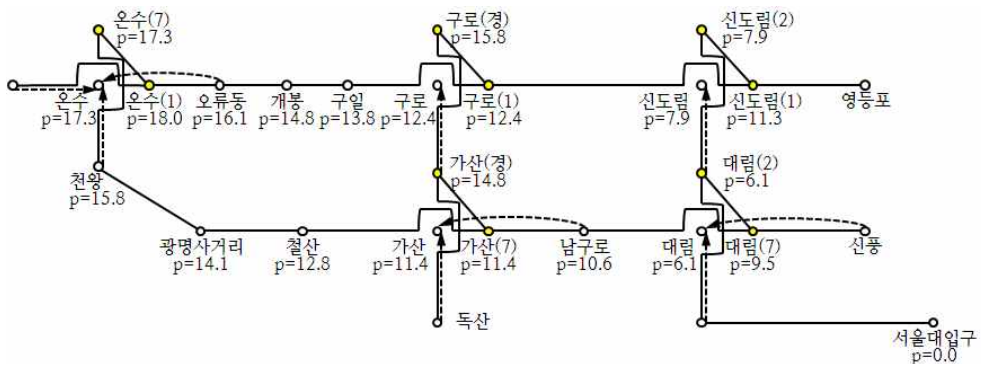
수정된 알고리즘은 다음과 같이 구성된다.



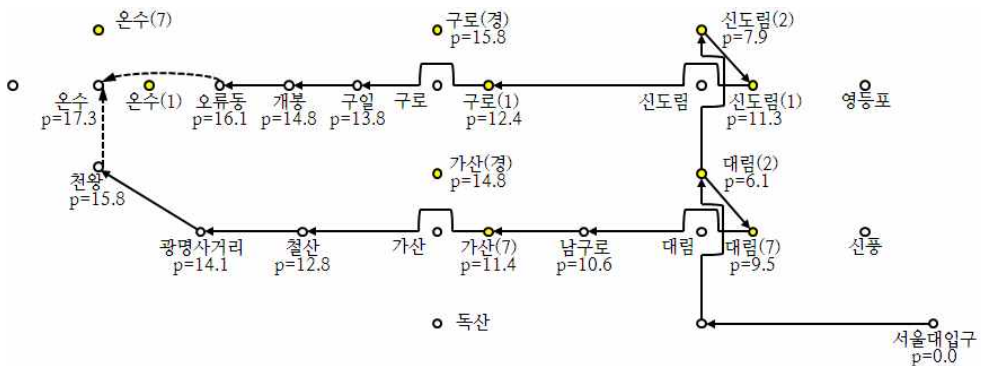


라) 효율적 경로 및 선택대안집합

경로기반 알고리즘은 무수히 많은 비합리적인 경로를 선택대안집합으로 고려해야 하지만 Dial 알고리즘은 효율적 경로라는 축소된 네트워크(reduced network)만을 고려하기 때문에 비합리적인 경로, 예컨대 루프형 태로 순환하는 경로(cyclic path), 지나치게 우회하는 경로 등에 배정되는 것을 사전에 방지할 수 있다. 실제로 Dial 알고리즘을 수도권 도시철도 네트워크에 적용했을 때 나타나는 효율적 경로는 실질적으로 이용될 가능성이 높은 경로만을 보여준다. 예를 들어, 서울대입구-온수 간 전체 네트워크와 효율적 경로는 각각 <그림 IV-17>, <그림 IV-18>과 같다.



<그림 IV-17> 서울대입구-온수 간 전체 네트워크



<그림 IV-18> 서울대입구-온수 간 효율적 경로

<표 IV-28> 서울대입구-온수 간 전체 경로와 효율적 경로

구분	경로	통행거리(km)			효율적 경로
		차내	환승	합계	
경로1	서울대입구-신도림-온수	14.6	3.4	18.0	○
경로2	서울대입구-대림-온수	13.9	3.4	17.3	○
경로3	서울대입구-대림-가산-구로-온수	16.0	6.8	22.8	×

<그림 IV-18>, <표 IV-28>에서 나타나듯이, 서울대입구-온수 간 효율적 경로는 전체 네트워크에서 가산(7)-구로(1)를 통과하는 경로3이 제거된 축소 네트워크로 나타난다. 경로3이 경로1이나 경로2보다 환승횟수가 많고 차내통행거리도 긴 점을 감안한다면 합리적인 결과로 판단된다. 즉, 효율적 경로의 정의는 실질적으로 이용될 가능성이 높은 경로만으로 선택대안집합을 구축한다. 한편, 다른 분석지점인 서울대입구-고속터미널 간, 서울대입구-도곡 간의 효율적 경로도 <표 IV-7>에서 제시한 실제 이용된 경로와 동일하게 나타난다.

마) 알고리즘의 정산

Dial 알고리즘은 음수가 아닌 값을 갖는 파라미터 θ 를 통해 이용자의 인지오차 수준과 경로의 선택확률을 정산할 수 있다. θ 값이 커질수록 이용자는 경로들의 통행시간 차이를 정확하게 인지하고 최단경로를 선택하게 되며, 반대로 θ 값이 작아질수록 이용자는 경로들의 통행시간 차이를 인지하지 못하여 대안경로를 동등한 확률로 선택한다. 알고리즘 결과의 정확성을 높이기 위해서는 이용자의 통행행태를 가장 잘 묘사할 수 있는 θ 값을 찾아야 한다.

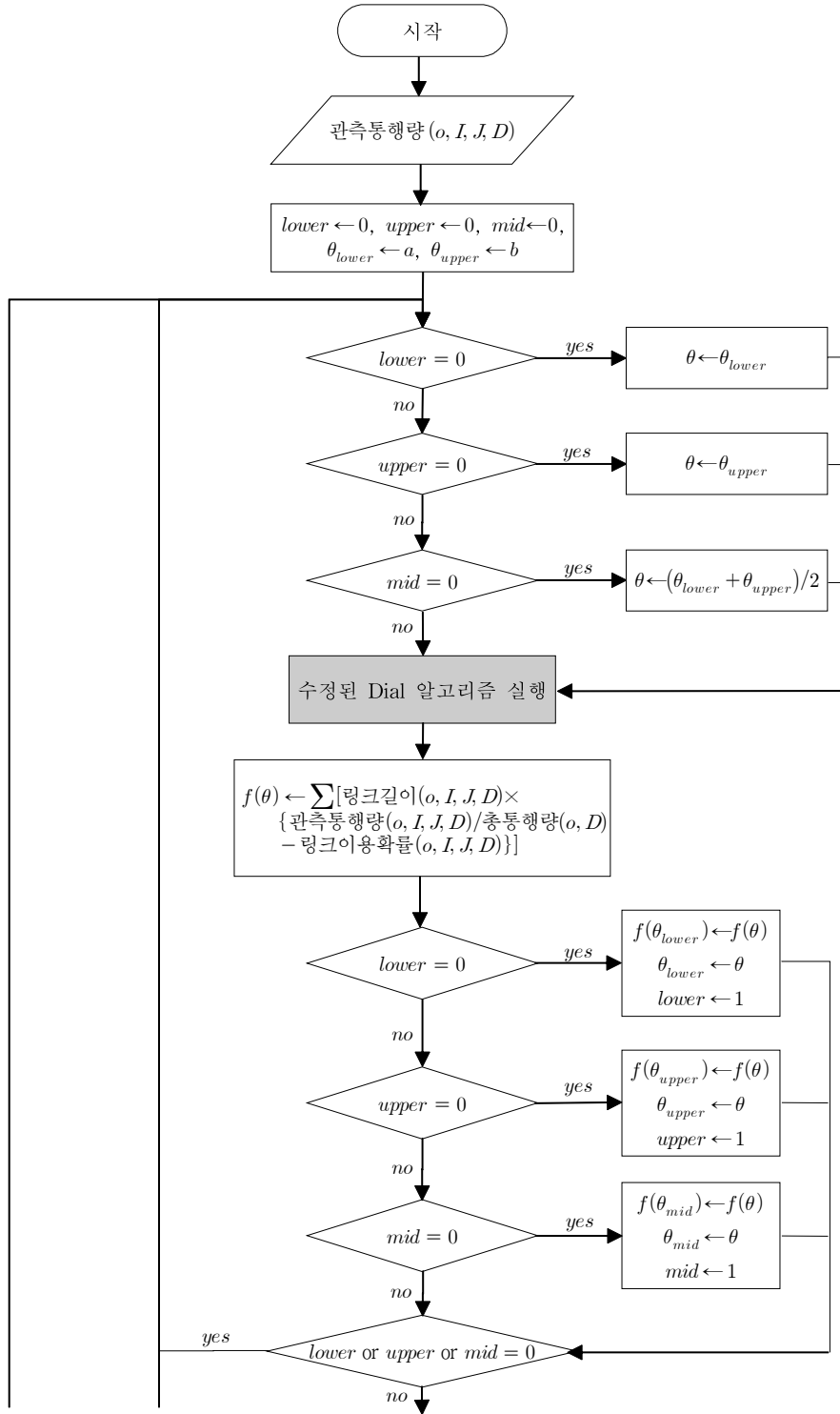
θ 값의 정산은 앞서 III장 4절에서 살펴본 Robillard(1974)의 방법을 준용하여 수행한다. Robillard(1974)는 최대우도 추정법을 통해 관측된 총통행거리와 예측된 총통행거리가 같게 되는 θ 의 최대우도 추정치 θ^* 를 구한다. θ^* 는 다음과 같이 $f(\theta)=0$ 를 만족하는 θ 의 해이다.

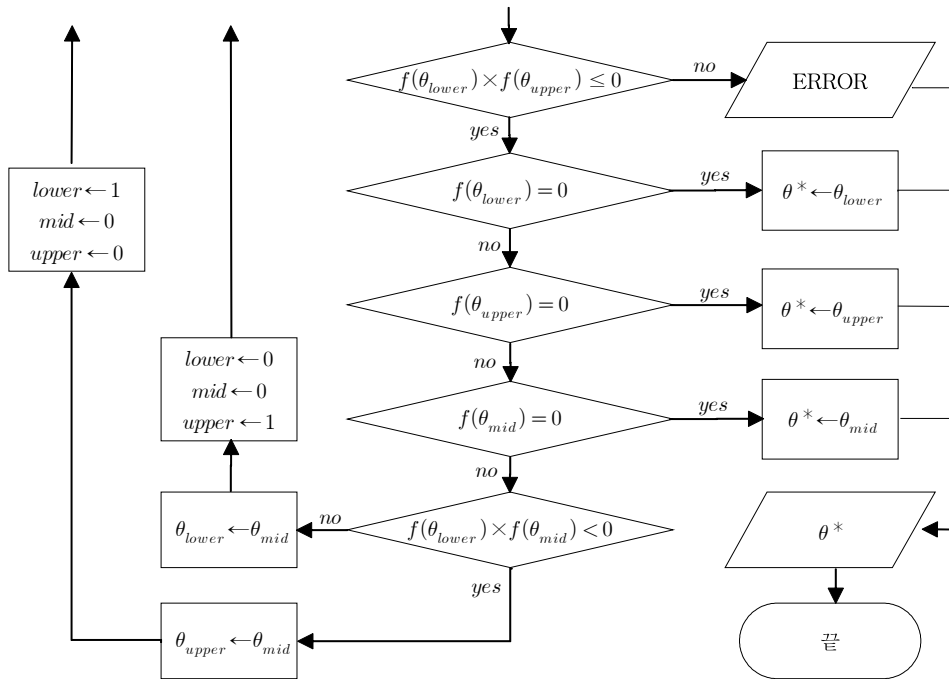
$$f(\theta) = \sum_{p \in P_{o,d}} x_p c(p) - T \sum_{p \in P_{o,d}} c(p) \text{Pr}(p) = 0$$

또는,

$$f(\theta) = \sum_{e \in E_{o,d}} x_e c(e) - T \sum_{e \in E_{o,d}} c(e) \text{Pr}(e) = 0.$$

위 식의 해 또는 근사값은 $f(\theta_1) \times f(\theta_2) \leq 0$ 인 θ_1 와 θ_2 의 값 사이에 존재하기 때문에 θ_1 과 θ_2 의 간격을 반복적으로 좁혀가는 방법을 이용하여 구할 수 있다. 본 연구에서는 출발역-도착역 별 링크의 이용확률을 구하기 위해 수정한 Dial 알고리즘과 정산 알고리즘을 결합하여 다음의 최종 알고리즘을 구축한다.





<그림 IV-19> 정산 알고리즘

위의 알고리즘을 적용하여 <표 IV-7>에서 제시한 출발지-도착지 쌍에 대해 θ 의 정산을 수행한다. 정산은 2009년 4월 데이터를 사용하여 수행한다.

<표 IV-29>는 서울대입구-도곡 쌍의 θ 를 정산하는 과정인데, $f(5) = 0.689$ 이고 $f(1.5) = -6.014$ 이므로 $f(\theta) = 0$ 인 θ 의 해는 1.5와 5사이에 존재한다. 앞서 설명하였듯이, <표 IV-29>에서 A경로의 61통행, B경로의 153통행은 ‘월별 지하철 여객 수송실적’ 자료에서 서울대입구-도곡(3), 서울대입구-도곡(분당선)의 일평균 통행량 값이며, 서울대입구-도곡 간 총통행량은 이 두 통행량의 합을 사용하였다. 서울대입구-도곡 구간은 환승지점이 같으므로, 환승링크의 길이는 $f(\theta)$ 결과에 무관하다. <표 IV-29>에서 환승링크 길이는 <표 IV-22>에서 구한 차내시간 기준 평균환승통행저항값인 3.4km를 적용하였다.

<표 IV-29> 서울대입구-도곡 구간의 θ 에 따른 $f(\theta)$ 의 결과

구간	관측 통행량 (a)	역간 거리 (b)	링크의 이용확률(=c)			d=a*b	e=b*c		
			$\theta=5$	$\theta=1.5$	$\theta=0.1$		$\theta=5$	$\theta=1.5$	$\theta=0.1$
서울대입구-낙성대	214	1	1.000	1.000	1.000	214	1.000	1.000	1.000
낙성대-사당(2)	214	1.7	1.000	1.000	1.000	363.8	1.700	1.700	1.700
사당(2)-방배	214	1.6	1.000	1.000	1.000	342.4	1.600	1.600	1.600
방배-서초	214	1.7	1.000	1.000	1.000	363.8	1.700	1.700	1.700
서초-교대(2)	214	0.7	1.000	1.000	1.000	149.8	0.700	0.700	0.700
교대(2)-강남	61	1.2	0.269	0.426	0.495	73.2	0.323	0.511	0.594
강남-역삼	61	0.8	0.269	0.426	0.495	48.8	0.215	0.340	0.396
역삼-선릉(2)	61	1.2	0.269	0.426	0.495	73.2	0.323	0.511	0.594
선릉(2)-선릉(분당)	61	3.4	0.269	0.426	0.495	207.4	0.914	1.447	1.683
선릉(분당)-한티	61	1	0.269	0.426	0.495	61	0.269	0.426	0.495
한티-도곡	61	0.7	0.269	0.426	0.495	42.7	0.188	0.298	0.347
교대(2)-교대(3)	153	3.4	0.731	0.574	0.505	520.2	2.486	1.953	1.717
교대(3)-남부터미널	153	0.9	0.731	0.574	0.505	137.7	0.658	0.517	0.454
남부터미널-양재	153	1.8	0.731	0.574	0.505	275.4	1.316	1.034	0.909
양재-매봉	153	1.2	0.731	0.574	0.505	183.6	0.877	0.689	0.606
매봉-도곡	153	0.8	0.731	0.574	0.505	122.4	0.585	0.460	0.404
합계						3179.4	14.854 (s)	14.885 (s)	14.899 (s)
(s)*총통행량							3178.71	3185.41	3188.39
$f(\theta)$							0.689	-6.014	-8.986

서울대입구-도곡 쌍처럼 대안경로의 환승횟수가 같을 경우, 환승가중치가 변하여도 대안경로 간의 총통행거리 차이는 변하지 않는다. 그러나 대안경로의 환승횟수가 다를 경우, 환승가중치가 변하면 경로 간의 총통행거리 차이가 변하기 때문에 θ^* 의 결과 역시 다르게 나타난다. 예를 들어, 대안경로의 환승횟수가 다른 서울대입구-고속터미널 쌍(<표 IV-7>참고)은 환승가중치에 따라 <표 IV-30>과 같이 θ^* 의 결과가 변한다.

<표 IV-30> 환승링크길이(가중치) 설정에 따른 θ^* 의 변화

환승 링크길이 (가중치)	2.7km (1.03)	2.8km (1.08)	2.9km (1.12)	3.0km (1.15)	3.1km (1.18)	3.2km (1.23)	3.3km (1.27)	3.4km (1.32)
θ^*	2.4520	2.2885	2.1455	2.0193	1.9071	1.8067	1.7164	1.6346

위의 결과는 미지의 환승가중치 중 어떤 값을 적용하더라도 θ 의 정산과정을 통해서 그 영향력을 조절할 수 있음을 시사한다. 달리 말하면, Dial 알고리즘은 경로특성, 출발지-도착지 특성, 혹은 기타특성(기상상태, 통행시각 등) 등에 의해 다양하게 나타날 수 있는 환승에 대한 이용자의 통행행태를 θ^* 값을 통해서 유연하게 조절할 수 있으며 통행배정 결과도 민감하게 달라지지 않는다. 따라서 Dial 알고리즘은 환승가중치 추정에 소요되는 시간 및 비용을 절감할 수 있다. 본 연구에서는 기존 연구의 환승가중치 평균값인 1.32를 적용하도록 한다. 즉, 네트워크에서 환승링크의 길이를 3.4km로 입력하고 정산 및 분석을 진행한다.

<표 IV-7>에서 제시한 3개의 출발지-도착지 쌍에 대하여 정산을 실시한 결과는 <표 IV-31>과 같다. θ_{all} 은 모든 출발지-도착지 쌍에 대하여 하나의 θ 를 정산한 결과이고 $\theta_{고속}$, $\theta_{도곡}$, $\theta_{운수}$ 는 각각의 출발지-도착지 쌍 별로 θ 를 정산한 결과이다.

<표 IV-31> 출발지-도착지 별 θ 의 정산 결과

반복	전체		서울대입구 -고속터미널		서울대입구-도곡		서울대입구-온수	
	θ_{all}	$f(\theta_{all})$	$\theta_{고속}$	$f(\theta_{고속})$	$\theta_{도곡}$	$f(\theta_{도곡})$	$\theta_{온수}$	$f(\theta_{온수})$
1	1.500000	-17.0547	1.500000	-17.1084	1.500000	-6.0139	1.5	6.0662
2	2.000000	32.6180	2.000000	28.8178	2.000000	-4.9762	1	2.7666
3	1.750000	13.4503	1.750000	11.4564	2.500000	-3.9587	0.5	-5.7084
4	1.625000	-0.0406	1.625000	-1.0811	3.000000	-2.9659	0.75	0.3961
5	1.687500	7.0967	1.687500	5.5754	3.500000	-2.0016	0.625	-1.6910
6	1.656250	3.6317	1.656250	2.3498	4.000000	-1.0691	0.6875	-0.5051
7	1.640625	1.8222	1.640625	0.6608	4.500000	-0.1714	0.71875	-0.0268
8	1.632813	0.8975	1.632813	-0.2034	5.000000	0.6893	0.734375	0.1908
9	1.628906	0.4301	1.636719	0.2303	4.750000	0.2637	0.726563	0.0836
10	1.626953	0.1952	1.634766	0.0139	4.625000	0.0474	0.722656	0.0288
11	1.625977	0.0774	1.633789	-0.0947	4.562500	-0.0617	0.720703	0.0011
12	1.625488	0.0184	1.634277	-0.0404	4.593750	-0.0071	0.719727	-0.0128
13	1.625244	-0.0111	1.634521	-0.0133	4.609375	0.0201	0.720215	-0.0058
14	1.625366	0.0036	1.634644	0.0003	4.601563	0.0065	0.720459	-0.0024
15	1.625305	-0.0037	1.634583	-0.0065	4.597656	-0.0003	0.720581	-0.0006
16	1.625336	-0.0001	1.634613	-0.0031	4.603516	0.0099	0.720642	0.0003
17	1.625351	0.0018	1.634628	-0.0014	4.600586	0.0048	0.720612	-0.0002
18	1.625343	0.0009	1.634636	-0.0005	4.599121	0.0023	0.720627	0.0000
19	1.625340	0.0004	1.634640	-0.0001	4.598389	0.0010	-	-
20	1.625338	0.0002	1.634642	0.0001	4.598022	0.0004	-	-
21	1.625337	0.0000	1.634641	0.0000	4.597839	0.0000	-	-

<표 IV-31>에서 확인할 수 있듯이, 각 θ 의 정산 결과는 $\theta_{all}^* = 1.625337$, $\theta_{고속}^* = 1.634641$, $\theta_{도곡}^* = 4.597839$, $\theta_{온수}^* = 0.720627$ 이다. 전체 출발지-도착지 쌍에 대하여 정산한 θ_{all}^* 의 결과는 서울대입구-고속터미널 쌍에 대하여 정산한 $\theta_{고속}^*$ 의 결과와 가장 근접한다. 서울대입구-고속터미널 쌍이 제시된 3개의 출발지-도착지 쌍 중에서 총통행량이 가장 많아 정산시 가장 큰 영향력을 미치기 때문이다. 게다가 서울대입구-고속터미널 쌍만

환승횟수가 다른 선택대안경로를 갖는데, 환승링크의 길이를 3.4km가 아닌 다른 값으로 적용한다면 θ_{all}^* 와 $\theta_{고속}^*$ 의 정산 결과는 달라진다.

서울대입구-도곡 쌍의 $\theta_{도곡}^*$ 값은 약 4.60으로 서울대입구-온수 쌍에서 $\theta_{온수}^*$ 값이 약 0.72로 정산된 것에 비해 비교적 큰 값으로 나타난다. 이러한 정산 결과는 서울대입구-도곡 쌍이 대안경로 간 총통행거리 차이는 0.2km이지만 대안경로의 관측통행량은 61통행, 153통행으로 상당한 차이인 반면, 서울대입구-온수 쌍은 대안경로 간 총통행거리 차이가 0.7km이지만, 두 경로의 관측통행량은 25통행, 38통행으로 차이가 작게 나타나기 때문이다. 요컨대 서울대입구-도곡 쌍에서는 0.2km만큼 짧은 최단경로의 선택확률을 높이기 위해서 θ^* 가 상대적으로 큰 값으로 나타나고, 서울대입구-온수 쌍에서는 최단경로와의 차이가 비교적 큰 0.7km이지만, 두 경로가 유사하게 이용되는 통행행태를 설명하기 위해 θ^* 가 작은 값으로 정산된다. 관측통행량 자료와 정산결과로부터 서울대입구-도곡 간 이용자가 서울대입구-온수 간 이용자에 비해 B경로(최단경로)의 의존도가 높다는 것을 알 수 있는데, 이러한 이용자의 경로선택 패턴은 단지 통행거리 변수만이 작용한 것은 아니라고 판단된다. 왜냐하면, 서울대입구-도곡 쌍에서 대안경로 간의 통행거리 차이가 다른 출발지-도착지 쌍의 대안경로 간의 통행거리 차이보다 작지만 경로 간의 관측통행량 비율차이는 크게 나타나기 때문이다.

서울대입구-도곡 쌍에서는 통행거리 차이 뿐만 아니라 출발지-도착지 간의 경로특성, 이를테면 A경로는 혼잡한 2호선 구간(교대~선릉)을 통과하기 때문에 차내의 혼잡도가 높다는 점, 또는 다른 분석구간보다 총통행거리가 짧으면 최단경로에의 의존도가 높다는 점 등이 경로선택에 영향을 미쳤을 가능성이 있다. 그런데 Dial 알고리즘은 앞서 <표 IV-30>에서 경로, 도착지, 통행자, 혹은 기타 특성에 의해 달라질 수 있는 환승가중치의 적용에 대해 논의하였던 것과 마찬가지로, 경로선택에 영향을 미치는 모든 변수를 찾아서 그 영향력을 구체적으로 추정할 필요없

이 θ^* 를 통해 이용자의 통행패턴을 묘사한다는 점을 알 수 있다. 이처럼 Dial 알고리즘은 설명하기 어려운 통행패턴이 나타나는 출발지-도착지 쌍에서도 파라미터 θ 의 정산을 통해 유연하게 분석할 수 있다는 장점을 갖는다.

바) 알고리즘의 적용

2009년 4월 통행량 자료로 정산한 각 출발지-도착지 쌍의 θ^* 값을 적용하여 링크의 이용확률과 배정통행량을 구한 결과는 <표 IV-32>, <표 IV-33>, <표 IV-34>와 같다.

<표 IV-32> θ^* 별 링크이용확률 및 배정통행량(고속터미널)

O-D	링크	관측 통행량	링크이용확률				배정통행량(2009. 4)			
			θ_{all}^*	$\theta_{고속}^*$	$\theta_{도곡}^*$	$\theta_{운수}^*$	θ_{all}^*	$\theta_{고속}^*$	$\theta_{도곡}^*$	$\theta_{운수}^*$
서울대입구 - 고속터미널	서울대입구-낙성대	831	1.000	1.000	1.000	1.000	831	831	831	831
	낙성대-사당(2)	831	1.000	1.000	1.000	1.000	831	831	831	831
	사당(2)-사당(4)	26	0.032	0.031	0.000	0.227	26	26	0	189
	사당(4)-이수(4)	26	0.032	0.031	0.000	0.227	26	26	0	189
	이수(4)-이수(7)	26	0.032	0.031	0.000	0.227	26	26	0	189
	이수(7)-내방	26	0.032	0.031	0.000	0.227	26	26	0	189
	내방-고속터미널	26	0.032	0.031	0.000	0.227	26	26	0	189
	사당(2)-방배	805	0.968	0.969	1.000	0.773	805	805	831	642
	방배-서초	805	0.968	0.969	1.000	0.773	805	805	831	642
	서초-교대(2)	805	0.968	0.969	1.000	0.773	805	805	831	642
	교대(2)-교대(3)	805	0.968	0.969	1.000	0.773	805	805	831	642
	교대(3)-고속터미널	805	0.968	0.969	1.000	0.773	805	805	831	642

<표 IV-33> θ^* 별 구간이용확률 및 배정통행량(도곡)

O-D	링크	관측 통행량	링크이용확률				배정통행량(2009. 4)			
			θ_{all}^*	$\theta_{고속}^*$	$\theta_{도곡}^*$	$\theta_{운수}^*$	θ_{all}^*	$\theta_{고속}^*$	$\theta_{도곡}^*$	$\theta_{운수}^*$
서울대입구 - 도곡	서울대입구-낙성대	214	1.000	1.000	1.000	1.000	214	214	214	214
	낙성대-사당(2)	214	1.000	1.000	1.000	1.000	214	214	214	214
	사당(2)-방배	214	1.000	1.000	1.000	1.000	214	214	214	214
	방배-서초	214	1.000	1.000	1.000	1.000	214	214	214	214
	서초-교대(2)	214	1.000	1.000	1.000	1.000	214	214	214	214
	교대(2)-강남	61	0.419	0.419	0.285	0.464	90	90	61	99
	강남-역삼	61	0.419	0.419	0.285	0.464	90	90	61	99
	역삼-선릉(2)	61	0.419	0.419	0.285	0.464	90	90	61	99
	선릉(2)-선릉(분당)	61	0.419	0.419	0.285	0.464	90	90	61	99
	선릉(분당)-한티	61	0.419	0.419	0.285	0.464	90	90	61	99
	한티-도곡	61	0.419	0.419	0.285	0.464	90	90	61	99
	교대(2)-교대(3)	153	0.581	0.581	0.715	0.536	124	124	153	115
	교대(3)-남부터미널	153	0.581	0.581	0.715	0.536	124	124	153	115
	남부터미널-양재	153	0.581	0.581	0.715	0.536	124	124	153	115
	양재-매봉	153	0.581	0.581	0.715	0.536	124	124	153	115
	매봉-도곡	153	0.581	0.581	0.715	0.536	124	124	153	115

<표 IV-34> θ^* 별 구간이용확률 및 배정통행량(온수)

O-D	링크	관측 통행량	링크이용확률				배정통행량(2009. 4)			
			θ_{all}^*	$\theta_{고속}^*$	$\theta_{도곡}^*$	$\theta_{온수}^*$	θ_{all}^*	$\theta_{고속}^*$	$\theta_{도곡}^*$	$\theta_{온수}^*$
서울대입구 - 온수	서울대입구-봉천	63	1.000	1.000	1.000	1.000	63	63	63	63
	봉천-신림	63	1.000	1.000	1.000	1.000	63	63	63	63
	신림-신대방	63	1.000	1.000	1.000	1.000	63	63	63	63
	신대방-구로디지털단지	63	1.000	1.000	1.000	1.000	63	63	63	63
	구로디지털단지-대림(2)	63	1.000	1.000	1.000	1.000	63	63	63	63
	대림(2)-신도림(2)	25	0.243	0.242	0.038	0.376	15	15	2	24
	신도림(2)-신도림(1)	25	0.243	0.242	0.038	0.376	15	15	2	24
	신도림(1)-구로(1)	25	0.243	0.242	0.038	0.376	15	15	2	24
	구로(1)-구일	25	0.243	0.242	0.038	0.376	15	15	2	24
	구일-개봉	25	0.243	0.242	0.038	0.376	15	15	2	24
	개봉-오류동	25	0.243	0.242	0.038	0.376	15	15	2	24
	오류동-온수	25	0.243	0.242	0.038	0.376	15	15	2	24
	대림(2)-대림(7)	38	0.757	0.758	0.962	0.624	48	48	61	39
	대림(7)-남구로	38	0.757	0.758	0.962	0.624	48	48	61	39
	남구로-가산디지털단지(7)	38	0.757	0.758	0.962	0.624	48	48	61	39
	가산디지털단지(7)-가 산디지털단지(1)	0	0.000	0.000	0.000	0.000	0	0	0	0
	가산디지털단지(1)-구로(1)	0	0.000	0.000	0.000	0.000	0	0	0	0
	가산디지털단지(7)-철산	38	0.757	0.758	0.962	0.624	48	48	61	39
	철산-광명사거리	38	0.757	0.758	0.962	0.624	48	48	61	39
	광명사거리-천왕	38	0.757	0.758	0.962	0.624	48	48	61	39
	천왕-온수	38	0.757	0.758	0.962	0.624	48	48	61	39

<표 IV-32>, <표 IV-33>, <표 IV-34>의 결과를 살펴보면, θ_{all}^* 을 적용하여 구한 배정통행량은 관측통행량과 오차가 크게 나타나는데, 통행특성이 다른 다수의 출발지-도착지 쌍에 대하여 하나의 θ 값을 적용했기 때문이다. 또한 다수의 출발지-도착지 쌍에 대하여 하나의 θ 값을 정산할 경우 θ^* 값은 통행량이 많은 출발지-도착지 쌍의 통행특성에 편중된다. 한편, 각 출발지-도착지 쌍에 대하여 정산한 θ 값을 적용하여 구한 배정통행량은 100%까지의 정확성을 보인다. 기본적으로 각 출발지-도착지

쌍에 대하여 θ 값을 정산하는 것이 가장 정확한 결과를 도출할 수 있지만, 실제 네트워크에는 상당히 많은 출발지-도착지 쌍이 존재하므로 출발지-도착지 쌍의 통행특성(총통행거리, 혼잡도, 환승지점의 위치, 네트워크 토폴로지 등) 유형에 따라 그룹화하여 θ 를 정산하는 것이 현실적인 대안일 수 있다. 만일 총통행거리의 수준에 따라 θ 값을 정산한다면 장거리일 때 대안경로 간 통행거리 차이와 단거리일 때 대안경로 간 통행거리 차이가 경로선택확률에 미치는 영향을 다르게 반영할 수 있을 것이다. 그러면 로짓모형이 오차항을 동일하게(identically) 분포한다는 가정으로 인해 갖는 한계점(<그림 III-4> 참고)을 완화할 수 있다.

사) 알고리즘의 검증

2009년 4월 자료로 정산한 θ^* 값을 적용하여 구한 링크의 이용확률을 경로의 선택확률로 다시 정리하면 <표 IV-35>와 같다. 이를 2010년 4월, 2011년 4월 자료에 적용하여 배정통행량과 오차율을 구하면 <표 IV-36>과 같다.

<표 IV-35> 경로선택확률

출발역	도착역	경로	경로선택확률	
			θ_{all}^*	개별 θ^*
서울대입구	고속터미널	경로A	0.032	0.031
		경로B	0.968	0.969
	도곡	경로A	0.419	0.285
		경로B	0.581	0.715
	온수	경로A	0.243	0.376
		경로B	0.757	0.624

<표 IV-36> 배정통행량과 오차율

도착역	경로	실제 통행량			배정통행량(오차율)					
					θ_{all}^*			개별 θ^*		
		'09.4.	'10.4.	'11.4.	'09.4.	'10.4.	'11.4.	'09.4.	'10.4.	'11.4.
고속 터미널	경로A	26	42	35	26 (0.02)	27 (-0.35)	27 (-0.23)	26 (0)	27 (-0.36)	26 (-0.24)
	경로B	805	812	811	805 (0)	827 (0.02)	819 (0.01)	805 (0)	827 (0.02)	820 (0.01)
도곡	경로A	61	61	59	90 (0.47)	91 (0.49)	87 (0.48)	61 (0)	62 (0.01)	59 (0)
	경로B	153	155	148	124 (-0.19)	125 (-0.19)	121 (-0.18)	153 (0)	154 (0)	149 (0)
온수	경로A	25	20	23	15 (-0.39)	18 (-0.09)	18 (-0.21)	24 (-0.05)	28 (0.41)	28 (0.23)
	경로B	38	55	52	48 (0.26)	57 (0.03)	57 (0.09)	39 (0.03)	47 (-0.15)	47 (-0.1)

Dial 알고리즘을 적용해 구한 경로의 선택확률은 앞서 살펴본 최적 전략 알고리즘, 경로기반 확률적 알고리즘에 비해 각 출발지-도착지 쌍의 통행패턴을 가장 잘 묘사하여 오차율이 작게 나타난다. Dial 알고리즘은 이용자의 경로선택에 대한 인지오차를 고려함으로써 단일 경로에 전량배정하는 문제점을 완화하고, 출발지-도착지 별로 특징적으로 나타날 수 있는 통행특성도 θ 의 정산을 통해서 유연하게 묘사할 수 있기 때문이다.

5) 소결론

지금까지 링크기반 결정적 알고리즘, 경로기반 확률적 알고리즘, 링크기반 확률적 알고리즘의 유형별 대표적인 알고리즘인 최적전략 알고리즘, 로짓기반 알고리즘, 본 연구를 통해 수정 및 보완된 Dial 알고리즘을 실제 수도권 도시철도 네트워크 중 경로 및 경로의 실제통행량을 합리적인 수준에서 판단할 수 있는 일부 출발지-도착지 쌍에 적용해 보았다. 각 알고리즘 별로 경로의 선택확률을 구하고 실제통행량과 배정통행량을 비교해 보았으며, 그 과정에서 세 알고리즘의 특성을 살펴보았다. 이를 정리하면 다음과 같다.

먼저, 링크기반 결정적 알고리즘인 최적전략 알고리즘은 대중교통 이용자의 통행행태를 설명하기 위하여 승차시간, 대기시간 인자, 대기시간 가중치, 승차시간 가중치, 접근시간 가중치 등 이용자 중심적인 다수의 변수와 파라미터를 사용하고 있지만 전량배정방식의 기본적인 가정때문에 이용자의 통행행태를 설명하는데 한계가 있다. 이용자가 노선을 선택할 확률을 노선의 운행빈도 비율로 계산함으로써 다른 서비스 특성보다 운행빈도에 더 큰 가중치를 두는 분석가 중심적인 알고리즘이라고 할 수 있다. 특히 경로선택 확률이 다른 서비스 수준은 고려하지 못한다는 점에서 배차간격(운행빈도의 역수)이 비슷하여 차이가 작고, 이용자의 경로선택에 영향을 미치는 요소가 많은 복잡한 수도권 도시철도 네트워크 분석에 부적합한 것으로 판단된다. 또한, 본 연구에 적용한 다수의 파라미터는 토론토 트랜짓 네트워크를 대상으로 추정된 값으로 분석의 시·공간적 범위가 다를 경우 범용적 적용이 가능한지에 대해 추가적인 검토가 수반되어야 한다.

경로기반 확률적 알고리즘은 경로선택 시 판단근거가 되는 통행자의 효용함수를 추정해야 하는데, 통행자의 경로선택에 영향을 미치는 속성변수의 종류와 그 영향력은 경로, 지역, 이용자, 분석 데이터, 혹은 기타(기상상태, 통행시각 등) 특성 등의 다양한 요인에 의해 달라질 수 있다.

통행특성이 유사한 출발지-도착지 별로 효용함수를 추정하기 위해서는 많은 수의 표본을 확보해야 하는데 이는 많은 시간과 비용이 수반된다. 현실적인 대안으로 제한적인 효용함수를 적용하지만, 효용함수의 속성변수만으로 설명이 되지 않는 특징적인 통행패턴이 나타나는 경로에 대해서는 오차가 발생할 수 있다. 게다가 속성변수의 설정과 파라미터 추정 결과에 따라 통행배정 결과가 민감하게 변할 수 있다. 경로기반 확률적 알고리즘은 효용함수의 한계점 뿐만 아니라, 분석단위를 경로로 한다는 점에서 경로나열문제가 나타난다. 출발지-도착지 간 무수히 많은 경로 중에서 실제 이용 가능성이 높은 경로만으로 선택대안집합을 구축하는 합리적인 방법론에 대한 논의가 필요하다.

반면에, 링크기반 확률적 알고리즘인 Dial 알고리즘은 위의 두 알고리즘이 갖는 문제점들을 극복한다. 우선, Dial 알고리즘은 이용자의 경로 선택에 대한 확률적인 가정을 전제하는 알고리즘으로서, 파라미터 θ 를 통해 인지오차를 조절할 수 있고 단일 경로에 전량배정하는 문제점을 완화한다. 개별 출발지-도착지 별로 정산된 θ 를 이용하면 높은 수준의 정확도로 통행행태를 묘사할 수 있다는 점도 확인하였다. 이용자의 경로선택에 대한 영향을 미치는 속성변수에 통행시간, 환승횟수, 환승시간, 차내시간, 대기시간, 이용자의 사회경제적 변수 등 상당히 많은 변수가 있으며, 근래에는 혼잡도, 네트워크의 토폴로지, 이용자에게 제공되는 토폴로지 정보 등도 유의한 변수로 보고되고 있다. 그리고 이들 변수들이 시·공간적 범위가 변함에 따라 영향력이 다르게 나타난다는 점을 감안한다면 파라미터 θ 만으로 통행자의 인지오차에 영향을 미치는 모든 변수의 영향력을 반영하고, 설명하기 어려운 이용자의 통행패턴도 묘사할 수 있다는 점은 Dial 알고리즘의 강점이다. 기본적으로는 각 출발지-도착지 쌍에 대하여 θ 값을 정산하는 것이 가장 정확하게 통행패턴을 묘사할 수 있겠지만, 대규모 네트워크에서는 상당히 많은 출발지-도착지 쌍이 존재하므로 출발지-도착지 쌍의 통행특성(총통행거리, 혼잡도, 환승지점의 위치, 네트워크 토폴로지 등) 유형에 따라 그룹화하여 θ 를 정산하는 것도

고려할 수 있다. 만일, 총통행거리를 기준으로 출발지-도착지 쌍을 그룹화하여 θ 를 정산한다면 오차항이 동일하게(identically) 분포한다는 가정에서 발생하는 로짓기반 알고리즘의 문제점도 완화할 수 있다.

한편, 로짓기반 알고리즘의 또 다른 문제점인, 효용의 오차항에 대한 독립성(independence)을 가정함으로서 발생할 수 있는 문제점은 도시철도 네트워크에서 찾아보기가 어려웠다. 도로 네트워크가 작은 변화로도 많은 경로수가 생성되는 복잡한 네트워크인 반면에, 수도권 도시철도 네트워크는 지정된 노선으로 운행하는 간단한 네트워크이고 환승역에서는 환승시간이 소요되므로 대안경로의 통행시간이 서로 상관성이 적게 나타나기 때문이다. 즉, 도로네트워크에서 주로 볼 수 있는 <그림 III-2>와 같은 구간은 찾아보기가 어렵다.

마지막으로, Dial 알고리즘은 링크와 노드 정보만을 이용하여 링크의 이용확률을 계산하는 링크기반 알고리즘이기 때문에 경로나열 문제가 발생하지 않는다. 게다가 효율적 경로의 정의는 비합리적인 경로, 예컨대 루프형태로 순환하는 경로, 지나치게 우회하는 경로 등을 사전에 효과적으로 제거할 수 있다. 분석지점에 대하여 실질적으로 이용된 경로만으로 축소된 네트워크(reduced network)를 구축함으로써 계산상으로 효율적이고 행태적으로 합리적인 통행배정 결과를 얻을 수 있다.

V. 결론

1. 연구의 요약 및 종합적 해석

현재 표준적으로 사용하고 있는 대중교통 통행배정 알고리즘은 링크기반 결정적 통행배정 알고리즘과 경로기반 확률적 통행배정 알고리즘이다. 그런데 두 알고리즘은 각각 통행자의 경로선택에 대한 ‘결정론적’인 가정과 분석단위가 ‘경로기반’이라는 측면에서 결함이 나타난다. 본 연구는 위의 두 알고리즘의 한계점을 극복하기 위한 대안으로 링크기반 확률적 통행배정 알고리즘을 제안하였다. 또한 링크기반 결정적 알고리즘, 경로기반 확률적 알고리즘, 링크기반 확률적 알고리즘의 각 유형별 대표적인 알고리즘인 최적전략 알고리즘, 로짓기반 알고리즘, 본 연구를 통해 수정된 Dial 알고리즘의 적합성을 비교, 분석함으로써 링크기반 확률적 알고리즘 적용에 대한 당위성을 논의하였다.

링크기반 결정적 알고리즘인 최적전략 알고리즘은 대중교통 이용자의 통행행태를 설명하기 위하여 승차시간, 대기시간 인자, 대기시간 가중치, 승차시간 가중치, 접근시간 가중치 등 이용자 중심적인 다수의 변수와 파라미터를 사용하고 있지만 전량배정방식이 전제하는 결정론적인 가정때문에 이용자의 통행행태를 설명하는데 한계가 있다. 다시 말하면, 최적전략 알고리즘에서 이용자가 노선을 선택할 확률은 효용극대화 이론을 기반으로하는 이용확률이 아닌 분석가의 관점에서 정의하는 결정론적인 측면의 이용 ‘비율’이라는 한계점을 갖는다. 특히 경로선택 확률이 노선의 운행빈도 비율의 함수로서 다른 서비스 수준은 고려하지 못한다는 점에서 배차간격이 비슷하여 차이가 작고, 이용자의 경로선택에 영향을 미치는 요소가 많은 복잡한 수도권 도시철도 네트워크 분석에 부적합한 것으로 나타났다. 이러한 한계점을 극복하지 못한 채, 다수의 변수를 사용하여 대중교통 이용자의 통행행태를 설명하려고 하지만, 오히려 다수의 파라미터를 추정하는데 많은 시간 및 비용이 소요되고, 분석의 시·공간

적 범위가 다를 경우 범용적으로 적용할 수 없어서 충분한 사전연구가 수반되어야만 한다는 어려움이 있었다.

경로기반 확률적 알고리즘은 통행자들이 같은 경로를 사용하더라도 그 통행시간을 다르게 인지한다는 ‘확률론적’인 가정을 전제함으로써 링크기반 결정적 알고리즘의 ‘결정론적’인 가정을 완화한다. 즉, 경로기반 확률적 알고리즘은 통행자 별로 각 경로에 대하여 효용을 갖고, 효용극대화의 원칙에 따라 가장 높은 효용을 갖는 경로를 선택한다고 가정한다. 일반적으로 효용은 관측이 가능한 고정효용과 관측이 불가능한 확률효용의 두 가지 요소로 구분된다. 여기서 확률효용에 대한 분포를 겔블(Gumbel)분포로 가정하면 로짓기반(logit-based) 알고리즘이 유도되고, 정규분포를 따른다고 가정하면 프로빗기반(probit-based) 알고리즘이 도출된다. 한편, 고정효용은 전통적으로 경로의 서비스 수준을 설명하는 변수, 통행자의 사회·경제적인 측면을 설명하는 변수 등을 사용하여 추정할 수 있다. 그러나 통행자의 경로선택에 영향을 미치는 속성변수의 종류와 그 영향력은 경로, 지역, 이용자, 분석 데이터, 혹은 기타(기상상태, 통행시각 등) 특성 등 다양한 요인에 의해 달라질 수 있으며, 새로운 변수들도 계속적으로 보고되고 있어서 효용함수의 추정은 시간 및 비용이 많이 소요되는 어려운 작업임을 알 수 있었다. 현실적인 대안으로 광범위한 범위에서 추정된 효용함수 또는 다른 연구의 효용함수를 빌려 적용하는데, 효용함수의 속성변수만으로 설명이 되지 않는 특징적인 통행패턴이 나타나는 경로에 대해서는 오차가 발생하였다. 게다가 속성변수의 설정과 파라미터 추정결과에 따라 통행배정 결과가 민감하게 달라지는 한계점을 살펴볼 수 있었다.

아울러, 경로기반 확률적 알고리즘은 분석단위가 ‘경로’이기 때문에 분석하는 출발지-도착지 사이에 존재하는 모든 경로의 정보가 필요하다. 실제 도시철도 네트워크와 같이 많은 수의 링크와 노드로 구성된 대규모 네트워크에서는 아무리 짧은 구간일지라도 비합리적인 경로를 모두 포함하여 무수히 많은 경로가 존재한다. 따라서 모든 경로를 나열(enumeration)

하는 것은 사실상 불가능하다. 간단한 예제 네트워크에서 간접적으로 살펴본 바와 같이, 경로기반 통행배정 알고리즘은 링크기반의 알고리즘보다 많은 수의 대안경로를 고려하며, 통행자들이 경로를 선택할 때 고려하지 않거나 실제로 이용될 가능성이 낮은 경로에도 선택확률을 배분함으로써 그만큼의 오차가 발생하였다. 한편, 경로나열문제를 해결하기 위한 노력으로 전체 경로집합에서 일부 경로만을 추출하는 방법론들이 제시되어 왔지만 경로 선택집합을 설정하는 기준이 타당한 근거없이 임의적이라는 점을 선행연구 고찰을 통해 알아보았다.

반면에, 링크기반 확률적 알고리즘인 Dial 알고리즘은 링크기반 결정적 알고리즘과 경로기반 확률적 알고리즘이 각각 통행자의 경로선택에 대한 ‘결정론적’인 가정을 전제한다는 점과 분석단위가 ‘경로기반’이라는 점에서 나타나는 문제점을 해결한다.

우선, 선행연구 고찰과 예제 네트워크를 기반으로 한 실증적 분석에서 Dial 알고리즘의 특성을 살펴보았다. Dial 알고리즘은 이용자의 경로선택에 대한 ‘확률론적’인 가정을 전제하는 알고리즘으로서, 파라미터 θ 를 통해 인지오차를 고려할 수 있고 단일 경로에 전량배정하는 문제점을 완화하였다. 또한 비합리적인 경로를 효율적 경로의 정의를 통해 사전에 제거하고, 링크와 노드 정보만을 이용하여 링크의 이용확률을 계산하는 링크기반 알고리즘이기 때문에 경로나열 문제가 없다.

수도권 도시철도 네트워크를 기반으로 한 알고리즘의 적합성 평가에서도 앞서 살펴본 Dial 알고리즘의 장점을 살펴볼 수 있었다. 이에 앞서, 기본적으로 승용차 통행배정을 목적으로 고안된 Dial 알고리즘을 대중교통 통행배정에 적용하기 위하여, 대중교통의 특징적인 통행특성인 환승통행을 고려할 수 있는 네트워크 표현방식을 제안하였다. 또한 전통적 Dial 알고리즘은 특정 출발지-도착지 쌍에 대한 경로분석이 용이하지 않는데 이러한 한계점을 극복하는 수정된 형태의 Dial 알고리즘을 구축하였다. 파라미터 θ 의 정산은 자료 상의 한계로 경로 및 경로의 관측통행량을 합리적으로 판단할 수 있는 일부 출발지-도착지 쌍에 대하여

Robillard(1974)의 정산방법을 준용하여 실시하였다. 정산 과정에서 실제 통행을 묘사하는 θ 의 개략적인 값과 Dial 알고리즘의 특성을 살펴볼 수 있었다. 예컨대, 수도권 도시철도 네트워크를 대상으로 개별 출발지-도착지 별로 정산된 θ 를 이용하면 높은 수준의 정확도로 통행행태를 묘사할 수 있었다. 이용자의 경로선택에 대한 영향을 미치는 속성변수에 통행시간, 환승횟수, 환승시간, 차내시간, 대기시간, 이용자의 사회경제적 변수 등 상당히 많은 변수가 있으며, 근래에는 혼잡도, 네트워크의 토폴로지, 이용자에게 제공되는 토폴로지 정보 등도 유의한 변수로 보고되고 있다. 그리고 시·공간적 범위가 변함에 따라 변수들이 영향력이 다르게 나타날 수 있다는 점을 감안한다면 파라미터 θ 만으로 통행자의 인지도 차에 영향을 미치는 모든 변수의 영향력을 반영하고, 설명하기 어려운 이용자의 통행패턴도 묘사할 수 있는 Dial 알고리즘은 다른 알고리즘에 비하여 유연하다는 사실을 알 수 있다. 이러한 장점 때문에 환승가중치 등을 명확하게 추정하지 않아도 θ 값을 통해 그 영향력을 조절할 수 있고, 통행배정 결과가 민감하게 변하지 않으며, 실제 오차율도 다른 알고리즘에 비하여 작게 나타났다. 파라미터 θ 는 각 출발지-도착지 쌍에 대하여 정산하는 것이 가장 정확하게 통행패턴을 묘사할 수 있겠지만, 대규모 네트워크에서는 상당히 많은 출발지-도착지 쌍이 존재하므로 출발지-도착지 쌍의 통행특성(총통행거리, 혼잡도, 환승지점의 위치, 네트워크 토폴로지 등) 유형에 따라 그룹화하여 θ 를 정산하는 것도 고려할 수 있다. 만일, 총통행거리를 기준으로 출발지-도착지 쌍을 그룹화하여 θ 를 정산한다면 오차항이 동일하게(identically) 분포한다는 가정에서 발생하는 로짓기반 알고리즘의 문제점도 완화할 수 있다.

한편, 로짓기반 알고리즘의 또 다른 문제점인, 효용의 오차항에 대한 독립성(independence)을 가정함으로써 발생할 수 있는 문제점은 도시철도 네트워크에서 찾아보기가 어려웠다. 도로 네트워크가 작은 변화로도 많은 경로수가 생성되는 복잡한 네트워크인 반면에, 수도권 도시철도 네트워크는 지정된 노선으로 운행하는 간단한 네트워크이고 환승역에서는

환승시간이 소요되므로 대안경로의 통행시간이 서로 상관성이 적게 나타나기 때문이다. 즉, 도로네트워크에서 주로 볼 수 있는 <그림 III-2>와 같은 구간은 찾아보기가 어렵다.

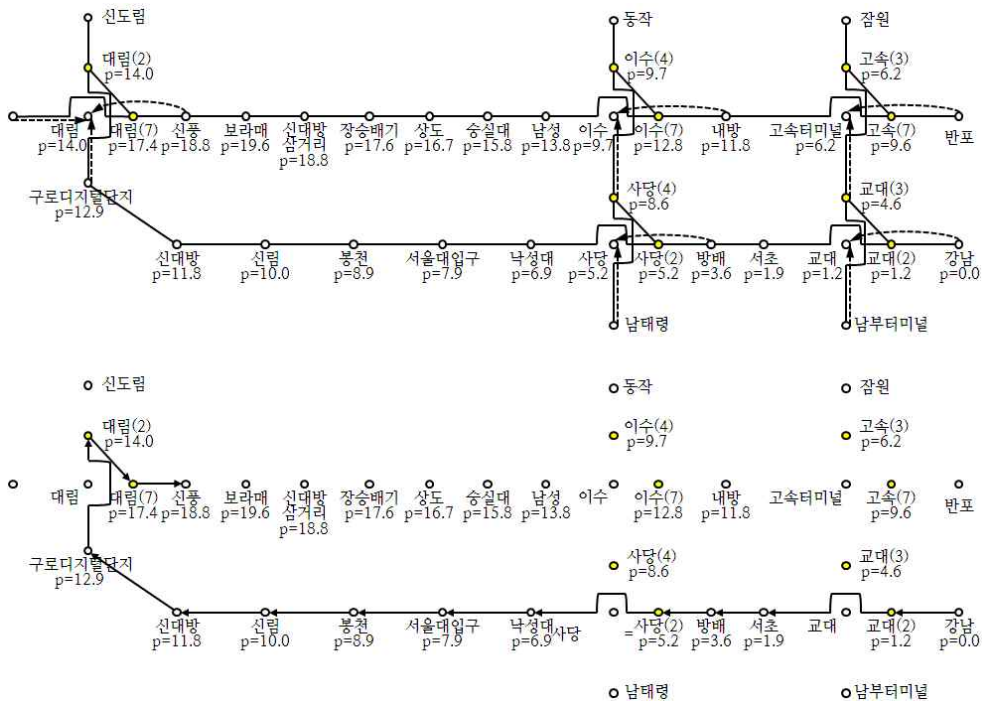
마지막으로, Dial 알고리즘은 링크와 노드 정보만을 이용하여 링크의 이용확률을 계산하는 링크기반 알고리즘이기 때문에 대규모의 수도권 지하철 네트워크에서도 경로나열 문제가 발생하지 않는다. 게다가 효율적 경로의 정의는 비합리적인 경로, 예컨대 루프형태로 순환하는 경로, 지나치게 우회하는 경로 등을 사전에 효과적으로 제거하였다. 분석지점에 대하여 실질적으로 이용된 경로만으로 축소된 네트워크(reduced network)를 구축함으로써 계산상으로 효율적이고 행태적으로 합리적인 통행배정 결과를 얻을 수 있었다.

요약하면, 링크기반의 확률적 통행배정 알고리즘인 Dial 알고리즘은 현재 대중교통 통행배정 알고리즘으로 표준적으로 사용하고 있는 링크기반 결정적 알고리즘과 경로기반 확률적 알고리즘의 내재적 결함을 해결한다. 또한, 많은 변수를 요구하지 않으면서도 실제 통행특성을 유연하게 묘사하고 이례적인 통행특성이 나타나는 구간에서도 범용적으로 적용할 수 있다는 장점이 있다.

2. 연구의 한계 및 향후 과제

링크기반 결정적 알고리즘, 경로기반 확률적 알고리즘, 링크기반 확률적 알고리즘을 실제 수도권 도시철도 네트워크에 적용하고 비교·분석 및 성과도를 평가하는 과정에서 알고리즘의 정확성을 판단하는 기준이 될 수 있는 경로의 관측통행량 자료가 부재하였다. 현재 수도권 도시철도의 운임 지불 시스템으로는 출발역과 도착역 사이의 각 경로(또는 링크)가 얼마나 이용되었는지 정확히 알 수 없기 때문이다. 본 연구에서는 이러한 자료 상의 한계점으로 인하여 경로 및 경로의 관측통행량에 대하여 약한 가정을 설정하고, 합리적인 수준에서 경로 및 경로의 관측통행량 판단할 수 있는 일부 출발지-도착지 쌍에 대해서만 분석을 실시하였다. 이러한 점 때문에 세 알고리즘 간의 상호 비교·분석만이 제한적으로 이루어졌으며 전체 출발지-도착지를 대상으로 도출된 통행배정 결과는 그 적합성에 대해 논의해 볼 수 없었다. 또한 Dial 알고리즘의 파라미터인 θ 의 정산 역시 소수의 출발지-도착지 쌍에 대해서만 이루어졌는데, 경로의 관측통행량 자료가 구축된다면 통행특성 별로 경로를 그룹화하여 정산하고 적용해볼 수 있을 것이다.

한편, Dial 알고리즘의 효율적 경로의 정의와 관련하여 흥미로운 점이 나타난다. 같은 두 지점을 출발지-도착지로 하더라도 그 방향에 따라 효율적 경로는 비대칭적인 형태로 나타날 수 있다는 점이다. 예를 들어, 수도권 도시철도 네트워크의 강남-신풍과 신풍-강남 구간은 효율적 경로가 비대칭적으로 나타난다. 먼저 강남에서 출발하여 신풍으로 가는 전체 네트워크 및 효율적 경로는 <그림 V-1>, <표 V-1>과 같다.

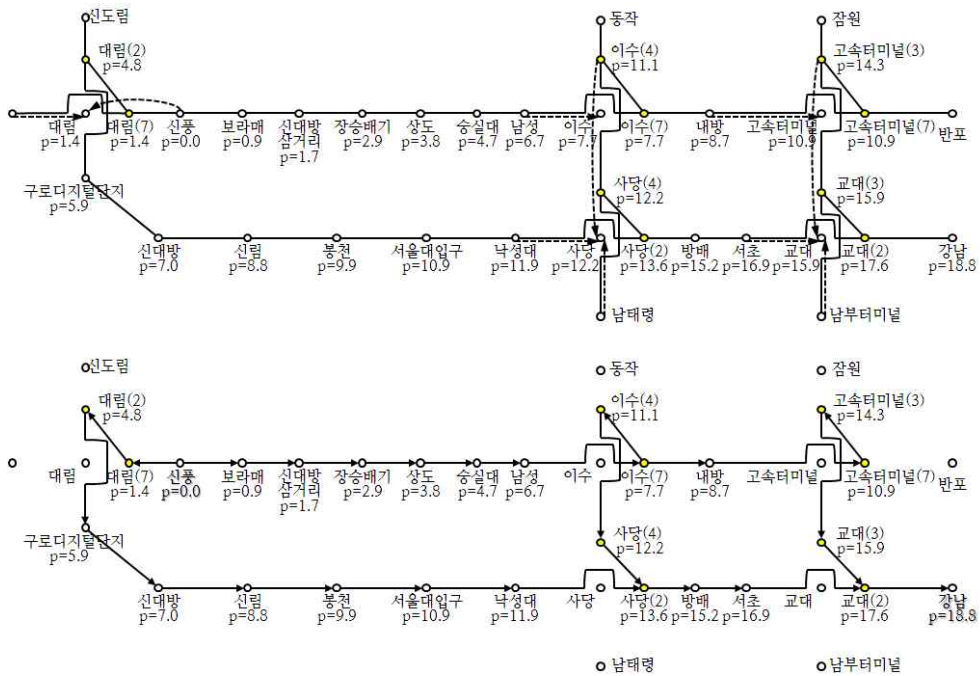


<그림 V-1> 강남-신풍 간 전체 네트워크와 효율적 경로

<표 V-1> 강남-신풍 간 효율적 경로

구분	경로	통행거리(km)			효율적 경로
		차내	환승	합계	
경로1	강남-대림-신풍	15.4	3.4	18.8	○
경로2	강남-사당-이수-신풍	14.0	6.8	20.8	×
경로3	강남-교대-고속터미널-신풍	13.7	6.8	20.5	×

<그림 V-1>과 <표 V-1>에서 확인할 수 있듯이, 강남-신풍 간에는 강남-대림-신풍으로 이어지는 경로1만이 효율적 경로로 나타난다. 반면에, 신풍-강남 간 효율적 경로는 <그림 V-2>, <표 V-2>에서 나타나듯이 강남-신풍 간에서의 결과와 차이가 있다.

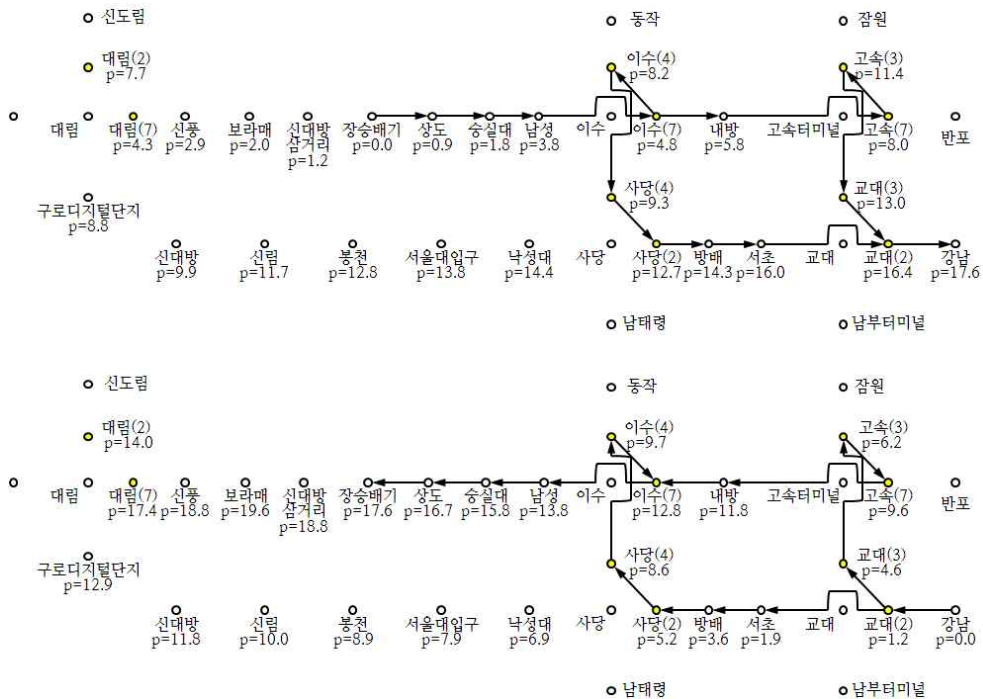


<그림 V-2> 신풍-강남 간 전체 네트워크와 효율적 경로

<표 V-2> 신풍-강남 간 효율적 경로

구분	경로	통행거리(km)			효율적 경로
		차내	환승	합계	
경로1	신풍-대림-강남	15.4	3.4	18.8	○
경로2	신풍-이수-사당-강남	14.0	6.8	20.8	○
경로3	신풍-고속터미널-교대-강남	13.7	6.8	20.5	○

강남-신풍 간에서는 경로1이 유일하게 효율적 경로로 나타난 반면에, 신풍-강남 간에서는 세 경로 모두 효율적 경로이다. 한편, 신풍-강남 (또는 강남-신풍)과 동일한 노선 및 환승지점이 존재하는 장승배기-강남은 <그림 V-3>, <표 V-3>에서 나타나듯이 강남-장승배기와 대칭적인 결과가 나타난다.



<그림 V-3> 장승배기-강남(위), 강남-장승배기(아래) 간 효율적 경로

<표 V-3> 장승배기-강남(강남-장승배기) 간 효율적 경로

구분	경로	통행거리(km)			효율적 경로
		차내	환승	합계	
경로1	장승배기-고속터미널-교대-강남 (강남-교대-고속터미널-장승배기)	10.8	6.8	17.6	○
경로2	장승배기-이수-사당-강남 (강남-사당-이수-장승배기)	11.1	6.8	17.9	○

출발지와 도착지에 위치에 따라 대칭적인 결과 또는 비대칭적인 결과가 나타나는 것은 효율적 경로의 정의가 네트워크의 형태(topology) 및 기하(geometry)를 반영하기 때문이다. 이처럼 효율적 경로의 정의는 네트워크의 형태(topology)를 반영함으로써 대중교통 이용자의 비대칭적 통행패턴을 설명할 수 있다는 가능성이 있다. 더불어, 앞서 논의한 바와

같이 효율적 경로의 정의는 지나치게 우회하는 경로를 사전에 효과적으로 제거하고, 실질적으로 이용될 가능성이 높은 경로만으로 축소된 네트워크(reduced network)를 구축함으로써 계산상으로 효율적이라는 점, 다소 엄격한 기준으로 통행배정되는 경로를 제한하여 일부 출발지-도착지 쌍에서 비합리적인 결과가 도출될 수도 있다는 점 등 다양한 특성이 나타난다. 게다가 효율적 경로의 정의는 경로기반 통행배정 알고리즘에서 경로나열문제를 해결하기 위해 사용하는 다소 임의적인 k -최단경로 알고리즘이나 유사경로 알고리즘 등을 대체할 수 있는 방안이 될 수 있다. 정리하면, 승용차 통행배정 알고리즘으로 고안된 Dial 알고리즘의 효율적 경로 정의는 대중교통 통행배정 분석에서 이용자의 비대칭적 경로선택을 설명할 수 있는 가능성과 다소 결정론적인 측면이 나타난다는 한계점을 동시에 갖고 있다. 이로부터, 효율적 경로 정의가 갖는 장점은 살리면서도 대중교통 이용자의 통행패턴에 부합하는 효율적 경로의 정의에 대한 연구가 심층적으로 이루어질 필요가 있겠다.

■ 참고문헌

- 국토해양부, 수도권 대중교통 승객의 열명 중 3명은 갈아탄다,
http://www.mltm.go.kr/USR/NEWS/m_71/dtl.jsp?lcmspage=1&id=155704805, 2010.(2012.6.30. 확인)
- 서울도시철도(2009), 「2009년 도시철도 수송계획」.
- _____ (2010), 「2010년 도시철도 수송계획」.
- _____ (2011), 「2011년 서울도시철도 수송계획」.
- 손상훈, 최기주, 유정훈(2007), “대중교통 통행배정을 위한 일반화비용 추정”, 대한교통학회지 25(2): 121-132.
- 신성일(2007), 「교통카드 Data를 활용한 대중교통 평가체계 구축 방안」, 서울시정개발연구원.
- 신성일, 노현수, 조종석(2005), “수도권 도시철도 수입금 정산 분석모형”, 대한교통학회지 23(5): 157-167.
- 신성일, 이창주, 김찬성(2010), “대중교통 카드(RF Card) 자료를 활용한 수도권 도시철도 운영기관 간 수입금 정산 방법론에 대한 연구”, 대한교통학회지 28(2): 7-19.
- 양창화, 손의영(2000), “서울시 지하철 이용객의 환승 관련 변수의 가치 추정 (선호의식(SP) 및 현시선호(RP) 분석을 이용)”, 대한교통학회지 18(4): 19-30.
- 이경재(2004), “환승역사의 동선체계를 고려한 환승패널티 추정-서울시 지하철 사례”, 서울대학교 석사학위논문.
- 임강원, 임용택(2003), 「교통망 분석론」, 서울대학교 출판부.
- 이재섭, 김익기(2001), “다이알 알고리즘을 이용한 다수단 대중교통 노선배정기법에 관한 연구”, 대한교통학회지 19(2): 53-60.
- 천승훈(2010), “교통카드 자료 기반 통합대중교통망의 확률적 통행배정모형

- 개발”, 서울대학교 박사학위논문.
- 한국개발연구원(2008), “도로·철도 부문 사업의 예비타당성조사 표준지침 수정·보완 연구(제5판)”.
- Akamatsu, T.(1996), “Cyclic flows, Markov process and stochastic traffic assignment”, *Transportation Research* 30B(5): 369-386.
- Azevedo, J. A., Santos Costa, M. E. O., Silvestre Madeira, J. J. E. R., and Martins, E. Q. V.(1993), “An algorithm for the ranking of shortestpaths”, *European Journal of Operational Research* 69: 97-106.
- Bekhor, S., Toledo, T., and Prashker, J.N.(2008), “Effects of choice set size and route choice models on path-based traffic assignment”, *Transportmetrica*, 4(2): 117-133.
- Bell, M. G. H.(1995), “Alternatives to Dial’s logit assignment algorithm”, *Transportation Research* 29B(4): 287-296.
- Caliper Corp.(2004) , “Travel Demand Modeling with TransCad”, 235-322.
- Daganzo, C.F., Sheffi, Y.(1982), “On stochastic models of traffic assignment”, *Transp. Sci.* 11: 253-274.
- Dial, R. B.(1971), “A probabilistic multipath traffic assignment model which obviates path enumeration”, *Transportation Research* 5: 83-111.
- Eppstein, D.(1999), “Finding the k shortest paths”, *SIAM Journal on Computing* 28(2): 652-673.
- Fisk, C.(1980), “Some developments in equilibrium traffic assignment”, *Transportation Research* 14B(3): 243-255.
- Floyd, R. W.(1962), “Algorithms 97, shortest path”, *Communications of*

- the ACM* 5(6): 345.
- INRO(1998), “EMME/2 User’s Manual”.
- Leurent, F., M.(1997), “Curbing the computational difficulty of the logit equilibrium assignment model”, *Transportation Research* 31B: 315–326.
- Liu, Y., Bunker, J., and Ferreira, L.(2010), “Transit Users’ Route-Choice Modelling in Transit Assignment: A Review”, *Transport Reviews* 30(6): 753–769.
- Maher, M.(1998), “Algorithms for logit-based stochastic user equilibrium assignment”, *Transportation Research* 32B: 539–549.
- Nielsen, O. A.(2000), “A stochastic transit assignment model considering differences in passengers utility functions”, *Transportation Research* 34B: 377–402.
- Nguyen, S., Pallottino, S., and Gendreau, M.(1998), “Implicit enumeration of Hyperpaths in a Logit Model for Transit Networks”, *Transportation Science* 32(1): 54–64.
- Parveen, M., Shalaby, A., and Wahba, Mohamed.(2007), “G-EMME/2: Automatic Calibration Tool of EMME/2 Transit Assignment Using Genetic Algorithms”, *Journal of Transportation Engineering* 133(10): 549–555.
- Raveau, S., Munoz, J. C., and de Grange, L.(2011), “A topological route choice model for metro”, *Transportation Research* 45A: 138–147.
- Robillard, P.(1974), “Calibration of Dial’s assignment method”, *Transportation Science* 8: 117–125.
- Sheffi, Y.(1985), “Urban transporatation networks: Equilibrium Analysis with Mathematical Programming Methods”, Prentice-Hall, New

York: 286-308.

- Si, B. F., Zhong, M., Zhang, H. Z., and Jin, W. L.(2010), “An improved Dial’s algorithm for logit-based traffic assignment within a directed acyclic network“, *Transportation Planning and Technology* 33(2): 123-137.
- Spiess, H., Florian, M.(1989), “Optimal strategies: A new assignment model for transit networks“, *Transportation Research* 23B: 83-102.
- Tobin, R. L.(1977), “An extension of dial’s algorithm utilizing a model of tripmakers’ perceptions“, *Transportation Research* 11: 337-342.
- Van Vliet, D.(1978), “Improved shortest path algorithms for transport networks“, *Transportation Research* 12: 7-20.
- _____(1981), “Selected node-pair analysis in Dial’s assignment algorithm“, *Transportation Research* 15B: 66-68.
- Yen, J. Y.(1971), “Finding the k shortest loopless paths in a network“, *Management Science* 17(11): 712-716.

Abstract

A Link-based Stochastic Algorithm for
Transit Assignment: Results from the
Urban Railway System in Seoul

Advised by

Prof. Chang, Justin Sueun

Submitted by

Jeong, Dongjae

August, 2012

Department of Environmental Planning
Graduate School of Environmental Studies
Seoul National University

Abstract

Transit assignment algorithm, as in the case of conventional vehicle traffic assignment algorithm classification, can be classified into deterministic assignment algorithm and stochastic assignment algorithm depending to variations in individual perceptions of passengers. In addition, it can be further classified into link-based assignment algorithm and path-based assignment algorithm depending on link or path unit of analysis.

The link-based deterministic algorithm and the path-based stochastic algorithm are standard transit assignment algorithms currently being used. However, these two algorithms have flaws in their respective aspects of the 'deterministic' supposition on the route choice of passengers and the path-based unit of analysis. This study proposes link-based stochastic algorithm as an alternative for overcoming the limitation of the two algorithms above. Dial's algorithm, which is the main representative link-based stochastic algorithm, will be applied to transit network to examine appropriateness of link-based stochastic algorithm. Prior to this, network representation will be proposed that can consider transfer, which is a particular characteristic of public transportation, since Dial's algorithm has been design for the purpose of vehicle traffic assignment. In addition, a revised algorithm will be developed that can solve the limitation of the

conventional Dial's algorithm that has the vulnerability in the path analysis according to a single-origin-destination pair. It will be conducted by applying the calibration method of Robillard(1974) for the calibration of parameter θ .

For the purpose of examining the feasibility of link-based deterministic algorithm, path-based stochastic algorithm and link-based stochastic algorithm, their respective representative algorithms - optimal strategy algorithm, logit-based algorithm and the revised Dial's algorithm of this study - will be applied example network and actual urban railway network in Seoul through which appropriateness of the application of link-based stochastic algorithm will be discussed.

First off, optimal strategy algorithm, which is a link-based deterministic algorithm, is using multiple passenger-oriented variables and parameters, such as boarding time, wait time factor, wait time weight, boarding time weight and auxiliary time weight, to explain the trip behavior of passengers, but it has a limitation in explaining the trip behavior of passengers because of its deterministic supposition all-or-nothing assignment premise. It was revealed to be unfit for analyzing complex urban railway network with optimal strategy algorithm that calculates path selection probability by function of service frequency because urban railway's headway is similar and it cannot consider other level of service. But there are many factors that affect the passengers' path selection.

Path-based stochastic algorithm alleviated the 'deterministic' supposition of link-based deterministic algorithm through its supposition that assumes that passengers perceive trip time differently even if they use the same path. However, estimation of utility function is a very difficult task that requires much time and cost because of many variables that affect passengers' path selection, and its limitation of the assignment result responding sensitively and changing according to the result was estimated. In addition, the issue of having to enumerate every path that exists between origin and destination arises in path-based stochastic algorithm since its unit of analysis is 'path'. Since countless paths exist between origin and destination in large network that consist of many links and nodes, it is, in fact, impossible to enumerate every path. Additionally, allocating probability also in paths that passengers do not consider when selecting path can cause errors.

On the other hand, Dial's algorithm, which is a link-based stochastic algorithm, overcomes the intrinsic flaw of link-based deterministic algorithm and path-based stochastic algorithm. It alleviated the issue of all-or-nothing assignment and reflecting variability in perceptions through parameter θ . In addition, it solved the path enumeration issue since it is a link-based algorithm that eliminates irrational paths through the definition of efficient paths and calculates the probability of use of link by using exclusively link and node information. Furthermore, its advantages were found in being able to flexibly describe trip

behavior of public transportation while not requiring many variables, as well as being able to extensively apply it also in the O-D pairs where unusual trip characteristics appear.

.....

Keywords : Transit assignment, Link-based deterministic assignment, Path-based stochastic assignment, Link-based stochastic assignment, Dial's algorithm

Student Number : 2010-23895